

9 September 1982

What kind of crisis for capitalism?

This week's meeting of the International Monetary Fund at Toronto will avoid an immediate crisis. But more international credit will not allow rich and poor to behave as if they live on different planets.

A few weeks ago, the International Monetary Fund put out a document that argued persuasively that inflation is an evil that the international community must somehow conspire to destroy. This week, in Toronto, the members of the fund will have arranged to do the opposite. Whether they agree that the amount of credit available internationally through the fund should be doubled (as the developing countries would prefer), increased by only a fifth (as the United States has been insisting) or will compromise somewhere in between is in a sense immaterial. What matters is that any increase in the money available to the fund will be potentially inflationary. Indeed, in present circumstances, any increase of the fund's ability to lend money will probably encourage inflation. So why have the members of the fund so quickly rejected the advice their organization was putting out in August? Because they have no choice.

The International Monetary Fund, one of the last creations of John Maynard Keynes, is not a bank in the strict sense but only a kind of bank. Its objectives were originally to deal with the problems that must arise from time to time in a world in which trade is free and when governments find themselves in financial difficulties. As with any other bank, the fund will then step in with a helpful loan, usually with strings attached (devaluation, fiscal probity and so on). The fund is not however a bank in the strict sense, because people (or governments) are not forever asking it to look after their money. Instead there is an arrangement under which the fund can call on its members to stump up a contribution if it thinks the time has come to lend. For the past quarter of a century, the system has worked well, justifying Keynes's advocacy of its necessity.

This week's decision in Toronto will break new ground in the history of the fund, and may even spell its end. A little recent history is relevant. When the price of oil was first drastically increased in the second half of 1973, the world's financial community seized on two immediate problems — the huge size of the financial surpluses that would accrue to the oil exporters and the plight of those developing countries that had neither oil nor the export capacity with which to buy it. In the event, the problems cancelled each other out: the oil-exporting countries lent their surplus funds to the commercial banks in the industrialized countries which then proceeded to lend on those funds to the countries that could not afford to buy oil, or the products that oil would make. Until a few months ago, this *ad hoc* system appeared to be working well. What has gone wrong is the discovery that many Western commercial banks have lent so much money to so many shaky creditors that some of them are in danger of going bust. The reasons why the members of the International Monetary Fund will have agreed this week on some increase of its lending capacity are simply that they cannot let their own commercial banks collapse under the weight of unpaid foreign debts. The snag, of course, is that neither the fund nor its member governments will be better able to control what happens now than the commercial banks have been.

The explanation is that the debtor countries have not been living as if to qualify for a loan from the International Monetary Fund. Rather, they have been borrowing from those commercial banks gullible enough to lend to them in the hope that they might avoid trouble for another few years — latterly, months or weeks. Yet every commercial branch manager knows that it is imprudent

to lend money to help a customer pay the rent, but that helping the customer, by capital investment, so as to be better able to do so is expedient and also socially beneficial. The International Monetary Fund is no longer in a position to require such a discipline from its creditors. First, its leverage as a quasi-bank has shrunk — in twenty years, its capacity to lend has shrunk by more than a third in relation to the monetary reserves of member states. Second, habituated as they are to living beyond their means the clients of the fund are neither willing nor able to cut back.

That is why this week's decision in Toronto cannot but be inflationary. Some clients of the fund will borrow money within the next eighteen months, will use it conveniently to repay their debts to the international commercial banks, but will not be able to promise that they will also promptly put their affairs in order. Who would reasonably ask of Mexico (see page 99) that it should deal quickly with these issues when its newly elected government cannot constitutionally take office until 1 December? And what is to happen to Argentina, unable to make a deal of any kind because the British Government has not been given proof that the Falkland Islands war is over? And then, worse still, what is to happen to those countries that have been buying oil for the past decade, but which have been unable to persuade banks of any kind of their credit-worthiness? Self-deprivation has allowed them to survive so far, but only passably. How far can self-deprivation be expected to go?

The Toronto meeting of the International Monetary Fund has frequently been described as a crisis of capitalism. It is nothing of the kind. Rather, it is the opposite — the occasion of the rediscovery of what nineteenth-century economists had to say about people's (or countries') interdependence. Chase Manhattan, or some other commercial bank, might be able to make Mexico (or some other country) bankrupt, but would then have to go to the wall itself. Middling governments such as those of Britain and West Germany may embark on policies intended to cure inflation domestically but then find themselves having to fuel it because they cannot stomach the consequences. Mitterrand's France may even have to acknowledge that its alliance of socialism and the printing-press cannot endure. And Reagan's United States may have to acknowledge that a \$100,000 million deficit is a recipe for beggaring its neighbours. So what the world needs is not a substitute for capitalism but some way of melding fiscal probity with an acknowledgement that neither the poor nor the rich can these days go it alone. Keynes, where are you now?

Paying for nuclear plants

Should US utility customers pay now or later for new nuclear plants? The utilities wrongly say now.

The substantial crop of nuclear power plants due to come into service in the United States in the next few years will have an unprecedented effect on the price paid by customers for electricity. Many customers are facing rate increases of 30 per cent; customers of one utility, Kansas Gas & Electric, will be charged 60 per cent more when the Wolfe Creek reactor comes into service in 1984. The utilities are seizing on these proposed rises as an opportunity to seek a major change in the way plant construction is financed. But the change they propose — charging



consumers for construction costs as they accrue, before the plant is in operation — is short-sighted and will benefit nobody.

Commercial utilities in the United States are subject to strict state regulation; in most states, the time-honoured principle of charging customers only for plant equipment that is "used and useful" is the rule. The utilities argue that the rule is outmoded. When nuclear plants used to cost a few hundred dollars per kilowatt and when interest rates were a few per cent a year, the addition of a new plant barely caused a ripple. Indeed, the savings in fuel costs would often outweigh the capital costs of a new nuclear plant, resulting in a decline in the price to customers. But costs have now reached staggering heights. A new nuclear plant, with a typical capacity of 1,000 megawatts, costs more than \$2,000 million, ten times as much as in the 1960s. High interest rates and construction delays add to the cost eventually charged to the consumer — whence the 30 per cent increases. The utilities want the right to charge consumers as they go. This "Construction Work in Progress" charge (CWIP) would be used to pay the dividends due to bond and stockholders during the period of construction. It would also result in a gradual phase-in of the price increase. Without CWIP, the utilities say, they have to borrow money to pay the bonds that become due during construction, and the extra cost of borrowing is ultimately passed on to the consumer once operation begins.

The utilities seem to have overlooked the principle on which the regulation of their charges is based — that present consumers should pay only for the services available to them now and not for services that will be available only to future cohorts of consumers. That principle, which underlies the regulation of all rates charged by public services in the United States, could not be overturned just like that to suit the present convenience of the electricity utilities. Moreover, to change the rules arbitrarily would be inequitable in quite tangible ways. For while it is true that by "paying now", a consumer would avoid having to pay for the interest charges accumulating during a construction project, he would then be denied the opportunity of investing the extra payments on his own account. "Pay now" costs at least as much as "pay later".

Although "pay now" seems to solve the utilities' immediate problems (not the least of which is the public relations problem of asking for 60 per cent increases), it glosses over a more fundamental difficulty. According to Professor Michael J. Driscoll of Massachusetts Institute of Technology, "the real solution is not to play around with CWIP, it's to get back to a reasonable construction schedule". Building nuclear plants used to take six years from start to finish; in France, they still do. But in the United States they now take 10 to 12 years. And the utilities' scapegoat — government delays in issuing licences — is only partly to blame. Driscoll points to inefficient management and the practice of custom-building each new reactor. France has seen the wisdom of that American innovation, mass production. But every delay increases the interest costs accumulated during construction, and the utilities would gain more by solving this underlying problem than by covering it up with CWIP charges.

Bluntly the burden of paying a 30 per cent or 60 per cent increase in prices is less for society as a whole than paying ahead of time for a nuclear plant that is not needed. The present system has the virtue of forcing the utilities to persuade bankers and investors that future demands will justify the next 1,000 megawatt power plant. The "pay now" policy, on the other hand, allows the utility to pass off some of the risk of its venture onto its consumers. And the track record of the industry in predicting future demand suggests that the consumers would be forced to make some bad investments. The industry has been notoriously optimistic — in 1974, the Edison Electric Institute (the commercial utilities' trade group) predicted a total US energy demand of 160 quads by the year 2000. By 1976, it was 140. In 1980, Exxon Corporation was saying 105. With the demand for electricity at a near standstill (rising only 0.3 per cent in 1981, in contrast with 7 per cent per year throughout the 1960s and early 1970s), the need for cautious planning is obvious. The marketplace should be allowed to continue to exercise a check on new power plant construction.

Press's forgotten recipe

Does the US scientific community consider Dr Frank Press's cure unpalatable?

Strange as it may seem, people have been ignoring Dr Frank Press, president of the US National Academy of Sciences. Press has had the temerity in the past six months to propose a long-term cure for the ills of US science. He has been observing that in these unprecedented times, federal funding on science has become part of the vulnerable discretionary federal budget and thus vicariously liable to cuts. What has happened this year, with the executive branch and Congress still haggling over money for the year that begins on 1 October, may be symptomatic of what is to come. So, Press says, the science community should negotiate a pact with government and industry to insulate science, to stabilize funding and to ensure that talented young scientists have opportunities.

In Press's five-point plan, government would agree to assured increases each year in the budget for basic research at a rate that would cover inflation and allow an extra 2 per cent of real growth. This "base programme" would be government policy. It would not rule out larger increases, but would ensure that overall funds did not fall below this level. Government should commit a further 1 per cent each year to "targets of opportunity", for research related to national needs or for new facilities. At the same time, industry would agree to double its present total contribution of \$50 million and would also start, with the government, a new fellowship programme. A central agency, perhaps the National Academy of Sciences, might serve as a collection and distribution agent, although the funds would be spent in fields where shortages are likely and selection could rest with individual agencies. A comparatively small amount of money, say \$10 million, would buy a thousand fellowships.

The nub is in the last of Press's five points. Scientists themselves, he says, should be able to find another 2 per cent by setting priorities within their own fields and by cutting out unproductive work. Then, the argument goes, universities could persuade government to ease regulations, such as the notorious Office of Management and Budget regulation on accountability, and let them keep the savings. Based on the estimated \$7,000 million the government spends on basic science and engineering, Press's potential cure could provide an extra \$694 million, an increase of almost 10 per cent, plus the use of up to \$114 million in funds that have been reprogrammed. But where should the extra money come from? Press suggests that the government should raid the development budget where, he believes, there is a lot of waste. (The Clinch River breeder reactor programme consumes about \$600 million each year, development of the B-1 bomber about \$700 million a year.)

The weakness of the plan is Press's assumption that scientists will centralize their lobbying efforts in Washington upon the single cause of a guaranteed overall increase. (Another is that Congress is all too familiar with special groups seeking to convert appropriations into entitlements.) Success for the plan would, however, bring predictability to the support of basic research and put scientists to work setting their own priorities, not leaving that to officials in Washington. And, of course, nothing in what Press has been saying suggests that the government has at this stage agreed that the scientific community could keep whatever savings would accrue from a more rational pattern of expenditure.

So why has the scientific community neglected the challenge? In the peer-ridden US system, changing tack in any direction is difficult. Arrogant specialization does not help. But Press's message is that if the research community cannot decide priorities for itself, the accountants and the lawyers will. Press seems to be echoing what Dr George Keyworth has been saying from the Old Executive Office Building next to the White House, but he has stronger credentials for being taken seriously. Indeed, with his experience of an earlier White House, Press is uniquely well placed to make what he has been saying tell. Why does nobody respond or even listen?

Norway's Arctic diplomatic fix

Poles' defection in Spitsbergen casts shadow

The two Polish scientists who defected last month from the Polish geophysical station at Hornsund (Spitsbergen) have created a delicate diplomatic problem for their Norwegian hosts and something of a crisis for Norwegian Arctic research.

Although sovereignty over Svalbard (the island group of which Spitsbergen forms part) was granted to Norway by international treaty in 1920, by the same treaty Norway undertook that the nationals of the signatory countries (of whom, with later accessions, there are now 41) should be treated on an equal footing with regard to economic activities.

The Soviet Union thus has two coal-mining bases, one near Barentsburg and another at Pyramiden. Similarly, under a memorandum from the Norwegian government to the other signatory states, foreign expeditions have free access to carry out scientific research in Svalbard as well as free access to the information on the islands held by the Norwegian Polar Research Institute.

The Polish research base at Hornsund is staffed by a team of 10 scientists from the Institute of Geophysics of the Polish Academy of Sciences. The research programme includes geological, geophysical, glaciological and meteorological work. A relief expedition, consisting of the 1982-83 scientific team plus a group of technicians to carry out routine maintenance work at the station sailed from Gdynia in July.

On 10 August, two of the scientists (it is not clear whether from the incoming or outgoing teams) decided to ask for political asylum, and radioed the Norwegian *Sysselman* (governor), announcing their intention to report in person to file their applications at the administrative base at Longyearbyen.

This transmission was, however, picked up by the Soviet Union's mining post at Barentsburg, which sent a helicopter to intercept the two Poles who were repeatedly "buzzed". One was driven back to base while the other fled into the hills and returned to Hornsund next day.

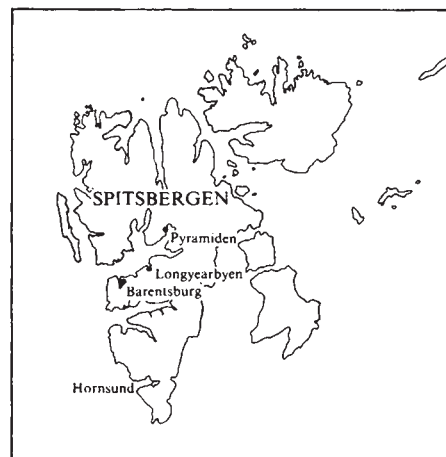
When the *Sysselman* heard what had happened, he sent out in his own helicopter to pick up the two would-be refugees; the Soviets, however, realized what was happening and sent their own helicopter to try and reach them first. In the words of one Norwegian official, *Sysselman* won the race, confirmed that the two men still wanted political asylum and took them back with him for questioning to Long-

yearbyen, whence they were dispatched to Oslo for "safe keeping". Their requests for asylum are now being processed; their identities are for the moment being kept confidential.

A number of diplomatic questions remain to be clarified. Military activity on the islands is strictly forbidden under the Svalbard Treaty of 1920; and it could be argued that the Soviet use of their helicopter in what can only be called a third-party police action could be construed as a breach of this agreement.

Furthermore, the Soviet helicopters are supposed to ply only between Barentsburg, Pyramiden and Longyearbyen, while some 44 per cent of Svalbard (including South Spitsbergen, where Hornsund is situated) comes under nature protection orders, where any technological encroachments, including motorized vehicles, are forbidden. Noise harassment of geese and other aquatic birds during the moulting season is particularly hazardous since it can scatter flocks and separate the young prematurely from their parents.

So far, the Norwegian Foreign Office seems concerned to keep the whole incident low-key. For their part, the Soviets can have little wish to fall foul of the Norwegian environmentalists. The Soviet Union extracts some 400,000 tonnes of coal



from Svalbard each year.

At Barentsburg, however, the reserves covered by the Soviet claims patent are almost exhausted, and output is now maintained only on the basis of a leasing agreement with the Norwegian mining company which owns the adjacent areas. A Norwegian Royal Decree of 1971 stipulates that no new mining, oil drilling or other industrial operations can be initiated without a thorough preliminary study of possible environmental consequences.

Vera Rich

Two Soviets banned from Perth

Canberra

The Australian government refused two Soviet scientists visas to attend the 12th International Congress of Biochemistry in Perth on 15-21 August, organized by the Australian Academy of Science under the aegis of the International Union of Biochemistry (IUB). The scientists are Professor Y.A. Ovchinnikov, vice-president of the Soviet Academy of Sciences who was to deliver a plenary lecture, and Professor S.S. Debov, also a member of the academy. Taking umbrage, the other 25 Soviet registrants boycotted the congress. It now appears that the International Council of Scientific Unions (ICSU) may ban Australia as a venue for international conferences at its meeting on 12 September.

The Soviets singled out for exclusion were regarded by the Australian embassy in Moscow as government men rather than scientists. Consequently the banning is consistent with government policy of not allowing high-ranking Soviets into the country. Following the Soviet invasion of Afghanistan in 1979, the government announced a package of sanctions against the Soviet Union, after a cabinet decision in January 1980. These moves included the suspension of a bilateral scientific and technological agreement signed in 1975, and the banning of mutual visits by

ministers and senior officials. A cooperative expedition on a Soviet research ship was called off, Soviet scientists in Mount Stromlo Observatory were sent packing, student scholarships to the Soviet Union were cancelled and visas refused to Soviet agricultural officials. In addition universities were advised to suspend academic exchanges but not multilateral conferences.

After these events, the Australian Academy of Science found itself in the invidious position of having to provide IUB with an assurance that no scientists would be barred. It sought from the government a "form of words" considered to be acceptable to IUB — "*bona fide* scientists from all countries will be admitted to Australia to attend the conference subject to the normal rules and requirements for visitor entry to Australia in operation at that time." But the then minister for immigration, Mr Ian MacPhee, wrote to the then foreign secretary of the academy, Professor Gordon Ada, in June 1980 — "I should stress that if this form of words is used, it must be understood by the academy that the normal rules in operation at the time of any conference might well exclude a particular scientist, a group of scientists or scientists of a particular nationality. I could not forecast what would be normal in say 12 months or 3 years time."

These reservations were not passed on to IUB. Unfortunately the possibility of Soviet scientists being banned as officials was not considered by the organizers at any time during the interim period. A spokesman from foreign affairs believes the problem could have been overcome by consultations beforehand.

The stage was thus set for the imbroglio at the conference, the first IUB congress in Australia. The Soviet participants waited until the last minute to apply for visas, and just a few days before the conference the ban was announced. Appeals from the president of the Australian academy (Professor Arthur Birch), the premier of Western Australia (Mr O'Connor) and the Swedish Academy were all to no avail. It was too late for the government to back off gracefully.

Reactions at the congress were mixed. Professor Frank Gibson of the Department of Biochemistry, John Curtin School of Medical Research, Australian National University, says "disappointment at not hearing Professor Ovchinnikov's lecture was the main reaction". Lack of travel funds, and the relatively high cost of internal air fares and accommodation limited the number of overseas registrants, particularly from the United States.

The academy now faces the unenviable task of persuading the government to change its policy in the week before the ICSU meeting. Failing this, it will try to delay any precipitate action at the meeting. Nevertheless international censure may be just the prod the government needs to review its policy on sanctions. Significantly, the International Union of Pure and Applied Chemistry succeeded in securing the admission of Israeli scientists into India for a conference on the applications of the Mossbauer effect held in Jaipur last December. **Vimala Sarma**

European space technology

Ariane set to go

The delay to the fifth launch of Ariane, rescheduled from last April to tomorrow, has had its effects on the vehicle's future launch programme. But careful re-organization of the launch schedule up until the end of 1984 seems to have ensured minimal damage to the launcher's commercial prospects.

The five month delay has been caused by problems with Marecs B, one of two satellites that the fifth Ariane will put into geostationary orbit. Marecs B had to be modified after problems with electrostatic charging on Marecs A, the sister satellite launched by the fourth Ariane late last year. The first effect of the delay has been to reduce the number of Ariane launchers this year from four to two. The next launch now scheduled for November will put Exosat, the X-ray observatory built by the European Space Agency, into polar elliptic orbit, more or less on its target date. But the launch of ECS 1, a communications satellite, is being delayed to early next year and Intelsat, the international telecommunications satellite organization, which was to have launched the sixth Intelsat V satellite aboard Ariane in October or December this year, has taken that part of its custom elsewhere.

Intelsat, however, has not reduced its overall commitment to Ariane. The launcher will now put Intelsat V numbers 7, 8 and 9 into orbit next year instead of numbers 6, 7 and 8. Next year's launch programme has been rescheduled to include six launches, the maximum possible with the present launch pad, instead of five originally planned.

Ariane is now fully booked with firm orders until the end of 1984, according to

Arianespace, the company that will take over commercial operation of the launcher some time next year. Just two places remain to be firmly negotiated for 1985 and a number of reservations, for which customers must pay \$100,000, are apparently in hand for future years. But Ariane has now lost some early reservations from customers who have eventually plumped for US facilities and Arianespace acknowledges that the fifth launch of the shuttle in November, which will put a satellite into geostationary transfer orbit, could take the edge off Ariane's seeming attractiveness.

The next few Ariane flights will still be formally in the hands of the European Space Agency. Arianespace will take over operations and start earning revenue from launches from the middle of next year, probably after the tenth flight. The company hopes that by 1985 the capacity of the launch facilities at Kourou in French Guiana will increase from 6 to 12 when the second launch pad, now under construction, comes into operation. Arianespace expects to start showing a modest but respectable profit in 1985. But future prospects will also depend on the success of the space shuttle. The latest estimate is that customers will have little to choose between the two on price. Reliability, service and convenience could be most significant. **Judy Redfearn**

Australian science politics

Labor's policy

Canberra

The Australian Labor Party, more hopeful than ever that it will form the next federal government but less certain now than a few weeks ago that the general election will come soon, formally adopted a new policy on science and technology at its biennial conference here in July. Although the policy is vague enough not to be a constraint on Labor members of parliament, the convention that conference decisions are binding on elected representatives may yet cause trouble.

The new science policy was adopted on the nod on the last day of the conference, when most delegates were packing up to go home. On the face of things, science and technology have only low priority, at least compared with the uranium issue. The conference heard a sustained and embittered debate on the question of whether Australia under Labor would honour existing uranium contracts. That issue remains unresolved, although the conference confirmed the principle that Labor will supply uranium abroad only on tough conditions.

The cornerstone of the Labor Party's policy on science and technology is economic — government spending will bring prosperity and economic growth. Mr William Hayden, the leader of the party, and his colleague with responsibility for

Now even MIT has troubles

Washington

Even well-off US universities are not immune from the financial squeeze affecting US higher education. The Massachusetts Institute of Technology (MIT), which has an operating budget of \$500 million, reports that it had a deficit of \$2 million for the year just ended. Next year's deficit could be \$4 million. Consequently, it has now announced plans to cut costs and, among other things, to lay off 400 employees.

Unlike both the California Institute of Technology and Stanford University, with which MIT is often compared, MIT has a relatively small endowment and thus a harder cushion for hard times. The current deficit is not MIT's first; it had a deficit in the early 1970s, when it lost the extra income and overhead funds that went with the Charles Stark Draper Laboratory, from which it severed its

connections at that time.

According to provost Francis E. Low, approximately two-thirds of MIT's operating budget is composed of funds that come in each year to operate facilities such as the Lincoln Laboratory, with little control or financial benefit to the rest of the institute. To control rising costs, then, the institute must look to the \$120 or \$130 million in general funds that pay for education and its administration. To prevent rising deficits, these activities will have to be reduced. There will be some faculty reduction through normal attrition, causing a loss of approximately 30 positions. Also, 200 other people will be allowed to leave through attrition in the next few years, he says, while an additional 200 will be laid off. Low thinks the institute can afford some belt-tightening: "It [MIT] functions in some respects in a luxurious way". **Deborah Shapley**

science and technology, Mr Barry Jones, have repeatedly said as much.

Perhaps significantly, the present Liberal-Country Party coalition government has responded by increasing support for the Industrial Research and Development Incentives scheme from A\$14 million to A\$50 million. More broadly, the government has taken to justifying a more interventionist approach to Australian industry by references to "market imperfections".

More animal care

Washington

Public pressure for a change in rules governing the care of laboratory animals has prompted the National Institutes of Health (NIH) to tighten up their requirements, whether or not Congress acts on the matter this year.

NIH will require that at least one non-scientist be appointed to each research institution's animal care committee — the groups charged with overseeing the treatment of experimental animals. This echoes one of the requirements in the bill now before Congress. (The bill, renumbered HR 6928, has cleared the House Science and Technology Committee and is now before the health and environment subcommittee of the Energy and Commerce Committee.)

NIH is also planning to mount a series of site visits to check on the extent to which researchers are complying with NIH guidelines. Under current rules, recipients of NIH grants must agree to adhere to the practices outlined in NIH's *Guide for the Care and Use of Laboratory Animals*. Failure to do so can result in loss of NIH funding. A major criticism of the system, however, is that too much reliance is placed on the good-faith assurances of the researchers and too little on inspection.

But NIH is not yet ready to order on its own another change in the proposed legislation: the mandatory accreditation of all research institutions by an independent group such as the American Association for Accreditation of Laboratory Animals Care (AAALAC), which sets standards and conducts routine inspections every three years. According to Dr William Raub, NIH's associate director for extramural research, "We think the concept of accreditation makes a lot of sense. But we're not yet, at least, willing to mandate it until we know more about the costs". Universities have claimed that bringing all NIH-supported institutions up to the AAALAC standards would cost \$500 million. The legislation before Congress would eventually require all institutions to meet the AAALAC standards, but would soften the blow by allowing ten years in which to meet this goal.

Stephen Budiansky

Labor is less inhibited about intervention. The July conference undertook that research and development expenditure should increase from the present 0.9 per cent of gross national product to a level "equivalent to that in other technologically advanced countries". Prudently, Labor did not specify a target figure.

What Labor offers is more pragmatic — a promise of a new division in the Department of Science and Technology to be concerned with industrial research, development and innovation. A Labor government would be a dutiful customer for new Australian products, and would back technical innovations with hard cash. It would also provide venture capital for new kinds of industrial developments, in biotechnology for example.

On issues such as these, the parties differ only moderately. Neither will increase

spending on basic research. The best hope is that Labor would not let research funds be eroded by inflation.

Long-standing disagreements persist, however, about the role of foreign corporations in high-technology industries. Labor would require foreign corporations to carry out some research and development in Australia, but Mr Barrie Jones seems well aware that he must find a form of words not so restrictive that foreign corporations will be driven away.

The new Labor policy also promises innovations to give high relief to Australian science. The conference approved a plan for an "Australia Prize" of A\$100,000 for humanitarian scientific achievement, to be administered by the Australian Academy of Science, and an international conference on Antarctic research to reinforce Australia's claim to the largest slice of the Antarctic.

Vimala Sarma

Mexican science in money trouble

Mexico City

In the face of Mexico's present economic crisis, the country's expensive and ambitious programme for science and technology now faces shocking sudden austerity. "A crisis can be good," says Pablo Rudomin, the distinguished neurophysiologist who is president of the Mexican Academy of Sciences. "It leads you to deep thoughts about where you're going. I think it [the crisis] will give us a sense of community. We'll have to work on priorities, and quality and on preserving what we have."

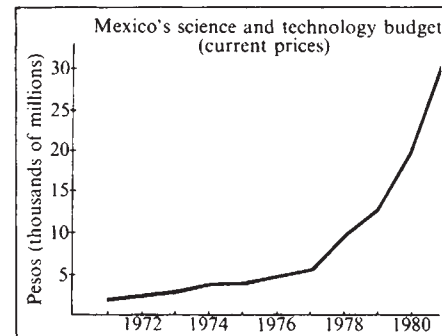
What has happened is that after a decade when the economy was flooded not just with domestic oil revenues but with loans from foreign banks (chiefly in the United States), Mexico has found that it cannot pay the instalments due on its debts, estimated at \$80,000 million. The payments were rescheduled in August, since when the Mexican banks have been nationalized. All sectors of the economy will have to be cut back, but just how will not be known until the new government of President Miguel de la Madrid Hurtado takes office on 1 December.

"Things will change, but probably for the good," echoes economist Manuel Gollás, who holds a key political post as head of the advisers to the director general of the Consejo Nacional de Ciencia y Tecnología (CONACYT), the Mexican science agency. One irony of the boom, he reflects, is that sometimes "we had more money than we did worthwhile projects".

Supported with government oil revenues and announced with a great fanfare at the beginning of President José López Portillo's 6-year term, Mexico's science and to technology programme was designed to strengthen the infrastructure of science and to encourage Mexican self-sufficiency in technical matters. The programme was nothing if not ambitious. Basic research did well, but there were also

special priority programmes in fields such as agriculture and nutrition. The energy programme included work on nuclear and solar energy production. There were schemes for supporting industrial research, and the government had emphasized the importance of training graduate students.

Interviewed in their respective offices in Mexico City, both Rudomin and Gollás acknowledged that the programme had been criticized as being only a shopping list of unrelated projects, reflecting the interests of individuals and government agencies. Indeed, its execution reflected this pattern. CONACYT, which designed



the original programme and serves as the president's science adviser, actually controlled only 11–12 per cent of the total science and technology budget. The government gave the rest directly to universities (the major one being the Universidad Nacional Autónoma de México, whose main campus is in Mexico City). Money for the Instituto Politécnico Nacional, for example, comes from the Ministry of Education. Private industry and foundations play almost no part in sponsoring Mexican science.

More than half of CONACYT's budget for 1981, estimated at 3,899 million pesos, went for scholarships, another 20–30 per cent for administration, public relations, and publications. Grants for basic science directly to investigators in fact accounted

for only 5 per cent of the budget, Rudomin estimates, or 109 million pesos. With other funds, basic research claimed about 20 per cent of the total CONACYT budget.

The science programme built on a major educational effort launched by the government of President Luis Echeverría in 1970 to stress education, including science and mathematics in the Mexican schools (*Nature* 280, p.101; 1979). According to Mexican officials, in 1968–69 Mexico had around 600 PhDs. By 1982, after more than 26,000 scholarships had been awarded, it claims 15,000 technically trained graduates, of which an estimated 6,000 hold a PhD.

Language a barrier

Mexico's ambition of achieving excellence in basic science is complicated by a double-edged language barrier: few scientists in the English-speaking world read Spanish scientific journals, while not many Mexican scientists know English well enough to publish in English-language journals. Even among Latin American countries, Mexico ranks low in the prorated number of its scientific publications in English-language journals, according to one recent survey. Many Mexicans feel it is "traitorous", as one of them put it, to publish in English. The Mexican science agency was sharply criticized when it decided that its fourth publication should be in English. (This is *RD Mexico*, a colour magazine that is suspending publication because of the current crisis.)

Science Citation Index reports that of the 3,068 journals in the index, only 13 are published in Latin America. Of these, 12 are published in Spanish and one is trilingual. Three of the 13 are published in Mexico. Latin America as a whole, therefore, contributes less to the index than does East Germany, which has 40 journals, or Austria, which has 24.

Mexican scientists complain that they are damned either way. If they publish in an English-language journal, their colleagues cannot read it and if they publish in a Spanish journal, their peers abroad will not. Yet English-language journals solve this dilemma for them, they allege, by discriminating against Latin American submissions. One scientist said that he has had an easier time getting his papers accepted by English-language journals when his co-authors have had Anglo-sounding names.

In any event, Mexico has a particular disadvantage on the language question because it did not experience the waves of immigration from European countries (except from Spain) that have admixed the populations of Brazil, Argentina and Chile. Deborah Shapley

One result is that most Mexican scientists are younger than Rudomin, who is in his mid-forties, and his predecessor as president of the Academy, engineer Daniel Reséndiz, who is 42. "Mexican science is not more than 40 years old," says Rudomin. "The first thing we have accomplished is to have trust in ourselves." Rudomin himself returned to work in Mexico after a stint at the Rockefeller University in New York. He has striven, he says, to make other Mexican scientists abroad come home to do good science.

However, the science programme did not succeed in achieving its goal that 1 per cent of the Mexican gross national product (GNP) should be devoted to science and technology. Science and technology have been between 0.38 and 0.47 per cent of the Mexican GNP for the past decade, whereas in developed countries the proportion is more than 2 per cent.

On the other hand, Mexico's GNP has grown so fast that the monetary gains for science and technology have been huge. The annual increases in government spending on science and technology averaged 34 per cent in the decade 1971–81.

The present crisis has changed all this. Sitting in his modest cinderblock office, Rudomin says wistfully of the CONACYT

basic science budget, "This year it was going to be 400 million pesos. Unfortunately, we are going to have a change in the slope. Our concern now is to keep what we have achieved."

Whether that will be possible is an open question. The ambitious science programme — and the financial policies that have got Mexico into trouble were the work of the Portillo government. By custom, the new government of president-elect Hurtado will be able to reshape science policy entirely: political appointees such as Gollás are getting ready to leave their jobs. "Our programme in science and technology was oriented as far as possible to the goals of the national development programmes, the whole national economy and society. That connection should be maintained, Gollás says.

Rudomin met Hurtado in August and is on a committee of scientists appointed to advise the new president before he takes office on 1 December. "When there's an economic crisis, they will try to solve the immediate problems," Rudomin says. "I think it depends a lot on us, whether we will be able to convince them that it is important to have this [science]. I think that it's going to be a difficult task."

Tabitha Powledge

Woods Hole laboratory

Summer camp seeks more funds

Woods Hole, Massachusetts

The passage of Labor Day last Monday will allow the biologists' favourite summer camp, the Marine Biological Laboratory, to settle down to its year-round pre-occupation: fund-raising. By the centenary of the laboratory in 1988, the plan is to have raised \$27 million, much of which will be spent on the rehabilitation of laboratories. There are also fond hopes of augmenting the endowment fund, now more or less a pittance at just over \$3 million.



From candle-making to administration

According to Dr Paul R. Gross, director of the laboratory since 1978, money began to dry up in the early 1970s. One of the painful discoveries since then is that the laboratory has no funds with which to cover the cost of maintaining buildings, but there are also ambitions to get rid of wet laboratories immediately above parts of the library (which needs also to be extended by 10,000 square feet) and to rehabilitate the housing in which the

summer campers camp.

These anxieties seem not to have depressed this summer's visitors, more than 1,000 altogether. Clam chowder has been bounteously consumed. People have acquired a tan that should last until Thanksgiving, and the beaches have been as full as ever in the afternoons, at least until the weather turned cold towards the end of August.

The popularity of Woods Hole may be, in the long run, its most enduring asset. Gross says that the competition for places on the seven summer student courses has been as brisk this year as in the past. Three-quarters of the successful applicants turn out to be graduate students; the remainder are an interesting mixture of advanced undergraduates and postdoctoral people.

The competition from more senior people to spend the summer at the bottom right-hand corner of Cape Cod is probably, however, more influential. Some senior people turn up to help teach courses; others move their research projects lock-stock-and-barrel, bringing their assistants with them. Prudent applications to grant-making agencies specify the importance of access to fresh squid or some such animal, for the laboratory will charge bench-fees, while the cost of accommodation will be extra.

Woods Hole, a frankly elitist establishment, is thus a continuation of academic life by other means — means by which a person's knowledge of how to organize a

clam-bake or to sail a boat may partly offset his ignorance of, say, the mechanisms by which nuclear RNA is processed. Some regular summer visitors have opted out, saying that they prefer their vacations straight. The regulars remark on the importance of each evening's social encounters.

Financially, the seasonality of Woods Hole must be a serious handicap, which is why some importance is attached to two recent innovations — student courses in January (this year there were four) and the continuing research programme (which at present keeps 95 professionals active all year round). Even so, the laboratory (whose operating budget is just over \$6 million, excluding the cost of peripatetic research projects) reckons that it recovers only 42 per cent of the cost of its facilities from those who use them, and that it is at present running on a deficit of about \$750,000 a year.

Part of the problem is constitutional. Since it was established in 1888 to help young women from Boston learn some science, the laboratory has been managed by its members, of whom there are more than 650, who in turn elect the 36-member board of trustees which for practical purposes determines policy.

The climate of opinion seems slow to change (and remains unwilling to commit too much of the site to year-round activities). Indeed, Paul Gross calls himself a "one-man glue-pot" whose job is to hold the place together. Two tangible proofs of



Paul R. Gross — one-man gluepot

his success are apparent — reluctant agreement that there should be a fund-raising effort of any kind and the conversion of the candle-house left over from the New England whaling industry into a handsome administration building.

The plight of the laboratory is thus desperate but not serious. (People tend quickly to say that the Oceanographic Institute, split off from the laboratory in 1930, has greater difficulties.) Visitors nevertheless cannot help but wonder whether the first word in the laboratory's title is still apt, given that "marine" featured in the title of only one of this summer's seven courses.

Biotechnology index

Biotechnology beats bull market

Reflecting the US stock market's sudden surge in August, *Nature's* monthly index of biotechnology company stocks jumped 18.1 per cent from the base of 100 at which it began last June (*Nature* 12 August, p.599).

The *Nature* Biotechnology Index at the close of August (see box) shows that biotechnology companies outperformed two leading indicators of overall US industrial performance during the same period. The 18 per cent rise between 25 June and 27 August was greater than either the rise in the Dow Jones Industrial Average, which rose 10 per cent, or the rise in the Standard & Poor 400 Index which rose 7 per cent, as the table shows.

Biotechnology and other US stocks

Date	<i>Nature</i> Biotechnology Index	Dow Jones Industrial Average	Standard & Poor's 400 Index
25 June	100.0	803.8	122.09
30 July	102.7	808.6	119.95
27 August	118.1	883.47	130.75

Both of these general industrial indicators declined in July, whereas the Biotechnology Index rose to 102.7 in July.

The biotechnology companies as a group also did better than Standard & Poor's index of 12 leading US drug companies, which rose by only 3 per cent during the same two-month period. However, the drug index rose 11 per cent between the end of July and the end of August.

The *Nature* Biotechnology Index is weighted according to the total market value of each company's outstanding shares. It includes 15 representative biotechnology companies publicly traded in the United States. Three of the companies

are based outside the United States.

Wall Street analysts who follow the emerging biotechnology industry were not surprised that the companies as a group performed well during the sudden rally at the end of August that saw trading at record volumes. "People had been looking for refuges in this market" said one analyst at Oppenheimer and Company. "They were looking for safe stocks, buying defensively, selecting only stocks that looked like they would have high growth". More traditional US companies, in the steel, automobile and chemical industries, are unlikely to be favoured until traders think that the decline in interest rates will enable those ailing industries to recover, she said.

Among the biotechnology stocks listed in the *Nature* index, the best performance was by well-established firms, such as A/B Fortia of Sweden, Novo Industri of Denmark, and Genentech of south San Francisco, California. But the stocks of many of the smaller companies did not move at all and some declined. Analysts explained this by saying that active traders tend to avoid very small companies because a purchase can control the overall price. The fortunes of the smaller companies are affected by internal changes.

For example, Bio Logicals, of Ottawa, Canada, saw the value of its 5.9 million shares of outstanding stock decline from \$3 to \$2 per share between July and August, partly perhaps because it has been without a chief executive officer since Robert Bender left in May. Such changes make a big difference to a small company's owners, but little to the market as a whole. "It's a volatile market", as one Wall Street analyst said.

Deborah Shapley

Nature index of biotechnology stocks

1982 high	1982 low	Company (Headquarters)	Close previous month	Close 27 Aug	Change
32 3/4	16 1/8	A.B. Fortia (Sweden)	25 3/4	32 3/4	+7
8	2	Bio Logicals (Canada)	3	2	-1
7	3 3/8	Bio-Response (US)	4 7/8	4	-7/8
14 1/8	8	Cetus (US)	9	8 7/8	-1/8
11	6 1/8	Collaborative Research (US)	8 1/8	8 1/8	0
21 7/8	14 3/4	Collagen (US)	17 3/4	17 1/8	-1/8
8 7/8	5 3/4	Damon (US)	6 1/4	7 7/8	+1 3/8
17 1/4	11 1/4	Enzo-Biochem (US)	11 1/4	12	+3/4
28	6 3/8	Flow General (US)	7 1/8	10 1/8	+3
37 3/4	26	Genentech (US)	30 3/4	33 3/4	+3
3 3/8	2 1/4	Genetic Systems (US)	2 3/4	3 1/8	+3/8
17 7/8	9 7/8	Hybritech (US)	13 3/4	14 3/8	+7/8
10 3/4	6 1/4	Molecular Genetics (US)	6 1/8	10 3/4*	+4 3/8
46 3/4	34 7/8	Novo Industri A/S (Den.)	39 7/8	46 3/4*	+6 5/8
12 3/8	8	Monoclonal Antibodies (US)	9 1/4	9 3/4	+1/2

The *Nature* Biotechnology Index for August 1982 stands at 118.1. Base of 100 as of 25 June 1982. See *Nature*, 12 August, p.599. Close of month prices at the close of business on the last Friday of the month. Where stocks are traded over the counter, the price quoted is the bid price. For stocks traded on the American and the New York Stock Exchanges, the price quoted is the transaction price. Data courtesy of E.F. Hutton, Inc.

*High or low for this calendar year.

CORRESPONDENCE

On dotty letters

SIR — It gives your readers, I suppose, some light relief when you publish dotty letters on improbable ways of achieving peace in our time. Grey Dimond's letter (*Nature* 15 July, p.220) on the merits of an exchange of marriageable young men and women in attractive uniforms between the Soviet Union and the West has, however, a sound basis despite the more fantastic embellishments.

In the past three years or so, roughly the period since the Soviet invasion of Afghanistan and the intensification of the cold war, the numbers of people visiting the Soviet Union from the United Kingdom have increased. So too have the numbers of Soviet citizens visiting Britain — largely in parties of 500 to 700 on cruise ships and staying for only three or four days. Though the British government signs every two years an agreement with the Soviet Union which includes a commitment to facilitate cultural, scientific and educational exchange, it sometimes obstructs exchange, and paradoxically this has been happening just when it is most needed. In addition to the severe restrictions on Soviet citizens wishing to travel abroad, those who manage to obtain external passports often have much difficulty in getting UK entry visas. Travel in either direction may also have been restricted by cutting to about half by a British government body early last year of the number of flights permitted each week between the United Kingdom and the Soviet Union — primarily for commercial reasons.

Exchanges are vital precisely because of our differences with the Soviet Union. Whether the aim of a visit be business, scientific exchange, tourism, diplomacy or nuclear disarmament, there is bound to be an educational aspect, an adjustment on either side, however slight, of perceptions and understanding of the cold war enemy, and this can only be for the good.

STEWART BRITTEN

London N6, UK

Japanese IQ

SIR — Regarding possible causes of Japanese children showing a significantly higher IQ than their US and European counterparts, Magyar (*Nature* 17 June, p.532) suggested that exposure to written Japanese might play a role. I should like to point out, however, that the spoken rather than the written form of Japanese language has been found to be responsible for a characteristic, epigenetic lateralization of auditory processing in the two cerebral hemispheres of native Japanese-speakers regardless of the ethnic identity^{1,2} (but do not be misled by ref.3 based on a complete misunderstanding of the subject). In this respect, both Koreans and Chinese exhibit the same type of lateralization as do Europeans whereas the Japanese and Polynesians belong to a minor group with a different type of lateralization. A hypothesis attempting to relate IQ differences to the first language spoken during childhood could be tested with populations in Taiwan and South Korea — both of which have recently experienced remarkable leaps in economic growth.

Bhargava (*Nature* 1 July, p.8), on the other hand, suspects that a quick democratization of

education accompanying economic improvement may have been responsible for the very high IQ of Japanese children. However, literacy among the Japanese had perhaps attained a level of 45 per cent for the male and 15 per cent for the female population by the mid-nineteenth century, and at the beginning of the twentieth century practically all children in Japan were attending primary school⁴. This would indicate that the Japanese were not very much behind the Europeans in terms of literacy, and that democratization of education in Japan was neither rapid nor very recent, as suspected by Bhargava. After the war, however, competition among the young Japanese for opportunities of higher education was very much intensified. Competition now starts very early — even at the entrance examination for "elite" kindergartens which will facilitate later competition for higher education in "good" universities promising a good income and a career of life-long commitment. It is therefore conceivable that fierce competition rather than democratization in education is a more likely cause of the higher IQ among Japanese children. The IQ of Japanese children who have attended, for a number of years, Japanese schools overseas (as in London or Sydney) — and in an atmosphere much less competitive than that in Japan — could well be used to test the point.

A. SIBATANI

CSIRO Molecular and Cellular Biology Unit,
Sydney, NSW, Australia

1. Tsunoda, T. *Nipponjin no Nō* (The Japanese Brain) (Taishukan, 1978).
2. Sibatani, A. *J. Social Biol. Struct.* 3, 255 (1980).
3. Miller, R.A. *Japan's Modern Myth* (Weatherhill, New York, 1982).
4. Reischauer, E.O. *The Japanese* (Harvard University Press, 1977).

Disarmament ideas

SIR — Among the many documents presented to the UN General Assembly at its Second Special Session on Disarmament (7 June–10 July) was one from the Medical Association for Prevention of War (MAPW) containing a completely new proposal. This calls for "a comprehensive ban on all methods and means of warfare specifically designed to kill or injure by inducing human disease, as distinct from causing mechanical and/or thermal effects". The starting point for the proposal is not the usual one of weapon-types, but the type of effect of weapons on people. The distinction between disease-inducing and mechanical or thermal effects has to be defined for the disarmament diplomats. This can be done in terms of pathology and, in lay terms, by the contrast of disease-inducing for living things only with the similarities of mechanical and thermal effects on animate and inanimate.

The MAPW proposal builds on existing agreements and negotiations on chemical, biological and toxin, and "radiological" weapons; and on agreed prohibitions on the release of "dangerous forces" from nuclear electrical generating stations, the use of environmental modification techniques for hostile purposes, and the use of conventional weapons that leave fragments in the body which are undetectable by X rays (so hindering their surgical removal and hence enhancing the risk of infection). The proposed comprehensive ban would close loopholes in

the existing agreements and extend the range of prohibitions — notably by including "neutron" weapons.

Moreover, the proposal opens up a new approach to the question of a comprehensive disarmament programme. The established approach is across-the-board disarmament of all varieties of military power. Modern versions (the United States' and the Soviet draft programmes of 1962 and the draft programme considered by the Special Session) envisage across-the-board comprehensive disarmament in three stages. This is like eating a cake in layers, and it is a proven failure.

The MAPW approach is like eating a mango. The whole of the skin ("disease-making") is peeled off first. Successive bites or cuts into the fruit can be flexible in position, size and depth. A comprehensive ban on thermal attacks on people might follow naturally from the ban on "disease-making". (If it is impermissible to inflict disease on one's enemies, one also ought not to burn them.) Reduction in the global total of the nuclear arsenals to below the theoretical level that poses a threat of species' genocide might be the next step. A prohibition on the development of any new weapons of mass destruction should be achieved without difficulty, as discussions on this subject are already in progress. Complete nuclear disarmament might have to await the introduction of objective principles for fixing ceilings on military manpower and expenditure. Current implications are that these should be proportional respectively to population size and gross national product (GNP). MAPW has found that this would have little overall effect on the present global maldistributions, which could be considerably improved, however, if the maximum permitted levels of military manpower and expenditure were made proportional to the *square roots* of populations and GNPs respectively.

The MAPW approach to comprehensive disarmament does not aspire to the ultimate UN objective of "general and complete disarmament". It leaves behind a hard core of national conventional military power. The hope must be that, having eaten the fruit, the world would decide that the only thing to do with the stone is to throw it away.

JEFFREY SEGALL

Medical Association for Prevention of War,
London NW2, UK

Units rule OK?

SIR — We view the recent letter by H.H. Rossi (*Nature* 22 July, p.320) with some disquiet. Are we witnessing the beginning of a trend towards the more widespread adoption of that Worrying Hardy Annual, The Nameless Or Unqualified Number In Today's Science (WHAT NO UNITS)?

While recognizing the practical value of bare figures for purely personal or even small group use, we believe that in all published work Units Nevertheless Incorporate Total Simplicity Regarding Unequivocal Literate Expressions Of Knowledge (UNITS RULE OK).

V.R. O'SULLIVAN
W.A. CURTIN
G.F. KAAR

Department of Anatomy,
University College,
Cork, Ireland

NEWS AND VIEWS

Kaposi's sarcoma: an oncologic looking glass

from Jerome E. Groopman and Michael S. Gottlieb

To paraphrase Lewis Carroll in *Alice in Wonderland*, Kaposi's sarcoma has become "curiouser and curiouser". It is unlikely that over a century ago when Dr Morris Kaposi¹ described patients with "multiple idiopathic pigmented sarcoma of skin" he would envision the disorder becoming an intriguing clinical puzzle that spans oncology, virology, epidemiology, public health, immunology and sexual behaviour. Indeed, the current epidemic of Kaposi's sarcoma and acquired immunodeficiency syndrome may provide important insights into the pathogenesis, prophylaxis and rational therapy of neoplasia.

In the spring of 1981 the Center for Disease Control (CDC) in Atlanta received the first of a number of reports of *Pneumocystis carinii* pneumonia occurring in young, previously healthy, homosexually-active men². Shortly thereafter, an unusual dermal malignancy, Kaposi's sarcoma (KS), was observed in similar hosts³. KS is a well recognized, rare endothelial neoplasm which generally occurs among elderly men of Mediterranean and Jewish ancestry. KS also occurs in an endemic form in certain areas of equatorial Africa. The common epidemiological link between Kaposi's sarcoma, opportunistic infection and the initial American cases was that of male homosexuality. The patients manifested a profound deficiency in cellular immunity, with anergy to recall antigens on skin testing, lymphopenia and a reversal of the normal ratio of helper to suppressor/cytotoxic T-lymphocytes^{4,5}. This ratio, although not specific for this disorder, has become a marker for the acquired immunodeficiency syndrome (AIDS). As of August 1982, there are over 500 documented cases of acquired immunodeficiency syndrome and/or Kaposi's sarcoma. There have been over 175 deaths due to this syndrome since its recognition, and the epidemic appears to be spreading⁶. In January 1982 about one case per day was reported to the CDC; since January 1982, two to three cases per day are being reported. It is estimated that KS and AIDS will account for 2–3 per cent of all deaths among men 25–40 years old in New York City, Los Angeles and San Francisco. There are now well documented cases occurring in 27 states in the United States as

well as 8 foreign countries. Ninety-five per cent of the affected persons are men; 85 per cent are either homosexual or bisexual although recently there has been an increased incidence among heterosexuals. The groups at risk for the development of Kaposi's sarcoma and AIDS now include homosexually-active men, intravenous drug users, female prostitutes, Haitian immigrants to the United States, and most recently, adult male haemophiliacs. There have been several intriguing epidemiological studies conducted by the CDC. In Los Angeles, 9 of 13 cases studied over the past year had had contact with another person who either had or later developed AIDS. It is of note that no lesbians have developed the syndrome to date. Fifty per cent of the cases come from the New York metropolitan area, about 25 per cent from California and about 7 per cent from Miami. The increasing incidence of the epidemic and its apparent transmission by sexual contact, drug apparatus and blood products strongly suggest a viral agent. Although use of various 'recreational drugs', particularly amyl and butyl nitrate, was initially believed to be important in immunodeficiency, this is now unlikely as many individuals with AIDS/KS have never used such drugs. There may be a genetic predisposition to the syndrome, as about 50 per cent of homosexually-active men in the New York area who developed Kaposi's sarcoma were positive for HLA-DR 5; nonetheless, the syndrome has clearly occurred among genetically, ethnically and socially diverse groups. Of greatest concern is the significant morbidity and mortality of the disease. Approximately 40 per cent of patients die per year with AIDS/KS; nearly 70 per cent of the patients who developed the syndrome in 1980 have died.

In the light of the occurrence of AIDS among adult male haemophiliacs, the United States government is now studying the donation and processing of blood products including hepatitis B immunoglobulin, factor VIII concentrates, cytomegalovirus (CMV) immunoglobulin,

and the new hepatitis B vaccine, all of which are frequently obtained from homosexually-active men. Recommendations from the Federal Drug Administration and the Bureau of Biologics are expected shortly.

The dimensions of this epidemic and the number of individuals at risk are unknown. Homosexually-active men who are asymptomatic have been studied and found to have a helper to suppressor-T cell ratio intermediate between that of homosexual controls and homosexually-active men with AIDS. How many of these asymptomatic men will go on to develop AIDS and/or Kaposi's sarcoma is not known. Nor is it known how the epidemic can be contained, aside from avoiding intravenous drugs and perhaps promiscuous sex. Recommendations to homosexually-active men concerning their sexual practices are difficult for the practising physician and involve complex ethical and civil rights issues.

Clinically, the opportunistic infections seen in men with AIDS are quite varied. Certain protozoa, such as *Cryptosporidium* and *Isospora*, have caused severe diarrhoeal syndromes; these are generally pathogens in livestock. In addition, certain atypical mycobacteria have caused not only severe, fatal pneumonias but a miliary pattern. Fatal fungal infections, particularly *Candida* and *Cryptococcus*, disseminated viral infections with herpes simplex, varicella zoster and adenovirus, as well as infection with toxoplasmosis, *Pneumocystis carinii* and *Legionella* have all been seen in AIDS. Therapy for many of these viral, protozoal and atypical mycobacterial infections is poor and most hosts quickly succumb.

Previous studies of endemic Kaposi's sarcoma in equatorial Africa have associated development of the neoplasm with cytomegalovirus infection⁷. A number of hypotheses have been advanced correlating the development of Kaposi's sarcoma with acquired immune deficiency syndrome. Repeated cytomegalovirus infection among homosexually-active men may 'wear down' the immune system and lead to cellular immunodeficiency. Yet why did this become apparent only during the past 2–3 years? Is there a new strain of cytomegalovirus? Although this is

Jerome E. Groopman is in the Division of Hematology-Oncology and Michael S. Gottlieb is in the Division of Clinical Immunology Allergy, UCLA School of Medicine, Center for the Health Sciences, Los Angeles, Ca 90024.

possible, preliminary studies on CMV isolates from affected hosts have not been particularly revealing. Is there a new virus? The CDC is in the process of processing sera and tissue from affected hosts and introducing filtrates into animals. Short of this, it would be difficult to identify a new virus. Because of the outbreak of AIDS among Haitian immigrants, attention has recently been turned to the Caribbean. There is preliminary information that AIDS and KS may be occurring in an epidemic fashion in Haiti. Because the

human T-cell leukaemia virus has a tropism for helper inducer T-lymphocytes, and because this retrovirus is endemic in the Caribbean, AIDS patients are being studied for HTLV by Robert Gallo of the NCI.

Additionally, Burkitt's lymphoma, cloacogenic carcinoma of the rectum and immunoblastic sarcoma appear to be occurring among American homosexual men at an increased frequency. Each of these cancers has been associated with DNA viruses of the herpes family. The epidemic of AIDS and KS affords a unique opportunity to study the relationship between immune status and the pathogenesis of certain forms of neoplasia. Considering the striking mortality, clues to aetiology that lead to rational therapy are needed. □

1. Kaposi, M. *Archs Derm. Syphilis* 4, 265 (1972).
2. Gottlieb, M.S. et al. in *Morbid. Mortal. Weekly Rep.* 30, 250 (1981).
3. Hymes, K.B. et al. *Lancet* ii, 598-600 (1981).
4. Gottlieb, M.S. et al. *New Engl. J. Med.* 305, 1425-1431 (1981).
5. Siegal, F.P. et al. *New Engl. J. Med.* 305, 1439-1443 (1981).
6. DeWys, W.D. et al. *Cancer Treat. Rep.* 66, 1387 (1982).
7. Boldogh, I. et al. *Int. J. Cancer* 28, 469 (1981).

The geoid and geodynamics

from Bradford H. Hager

THE direct radar measurements of sea surface height from satellite GEOS 3 reported in this issue of *Nature* (p.117) by Ki-Iti Horai show that it is now possible to measure small changes in the shape of the geoid — the equipotential surface corresponding to mean sea level — with an accuracy of ± 50 cm. Detailed maps of geoid elevation with contours spaced as close as 1m can thus be produced (see Fig. 1a, p.118) and clearly show the effect of tectonic features such as deep sea trenches, outer arc rises, and oceanic ridges and troughs on the height of the geoid.

Variations in the height of the geoid are caused by horizontal variations in the density structure of the Earth. Early interpreters, assuming a rigid Earth, used the observed fluctuations of the geoid to bound the size and depth of the density contrasts that brought them about. Constraints were thus placed on the (supposed finite) strength of the Earth¹. Now, observations of the postglacial rebound and plate motions make it clear that the Earth interior behaves as a fluid, not a rigid material, over geological time, requiring interpretation of geoid anomalies in a dynamic system²⁻⁴. As is well seen in Horai's paper, the study and interpretation of the figure of the Earth is now moving from the static domain of physical geodesy to the domain of geodynamics.

Analysis of geoid anomalies in terms of density contrasts has always been hindered by the lack of a unique interpretation — an infinite family of distributions of density contrasts within a given body can satisfy

any observed gravitational potentials outside that body. Adding the extra constraint that the equations of motion in the fluid interior of the Earth must be satisfied has only recently been exploited in interpretation of the Earth's gravity field, but imposes a limit on the class of acceptable solutions.

The basic physics of the problem are relatively simple, but contain some subtleties. Consider a density contrast of spherical harmonic degree l placed at a radius r in a fluid mantle. The gravitational potential at the surface ($r = r_c$) is reduced by a factor $(r/r_c)^{l+2}$. The density contrast causes flow which leads to a downward deflection of the surface of the Earth above high density regions and upward deflection above low density regions. Trenches and oceanic ridges are the directly observable result of this type of deflection. Similar deflections occur at the core-mantle boundary and at any other internal layer boundaries which might exist. (Settling the controversial question of the existence of these boundaries⁵⁻⁷ may soon be possible using the techniques outlined here.) Solving the equations of motion within a fluid mantle⁸ shows that the deflection of the surface of the Earth is also approximately given as $(r/r_c)^{l+2}$. The gravitational effects of the surface deformation and the driving density contrast are of comparable magnitude and opposite sign and nearly cancel.

The gravitational field of the Earth thus provides the most sensitive of all experiments, the null experiment, where the net result is a small number determined by the difference of two large effects. The sign of the result depends on which of the effects is dominant.

As an example, consider the geoidal

signal resulting from the subduction of a cold, dense slab of lithosphere. If most of the stress from the sinking slab were directly transmitted to the outer surface of the Earth, the gravitational effect of surface deformation would be of greater magnitude than that from the slab itself, resulting in a net negative geoid anomaly. If the stress were transmitted more efficiently to a discontinuity deeper in the mantle, the dense slab, being closer to the surface, would dominate the gravity field, leading to a positive geoid anomaly.

Horai's geoid map shows that except for the localized area comprising the trench, the gravitational attraction of the slab is larger than that of the deformed upper surface. This shows that the weight of the slab is carried mostly by support from below such as would be provided by an increase of viscosity with depth or by the presence of a chemically induced density contrast between the upper and lower mantle. Fault plane solutions showing in general down dip compression in deep slabs are consistent with either of these interpretations⁹. An increase in effective viscosity with depth might be the result of stress-dependent rheology, as Horai suggests, or of the competing effects of increased pressure and temperature on rheology.

Many of the features of the global geoid are associated with subduction zones¹⁰, but there are many features such as the geoid lows over Siberia, Hudson Bay, and the Ross Sea, and the high over Africa, not associated with subduction. In these areas, we have no *a priori* knowledge of the sign of the causative density contrasts. They may well be the result of density contrasts not associated with present plate motions¹¹. In particular, Anderson¹² has recently suggested that the geoid highs correspond to anomalously hot mantle, with geoid lows corresponding to colder mantle. This scenario is plausible if in these regions the gravitational effects of surface deformation are large, consistent with efficient coupling of stresses to the upper surface.

How can such a hypothesis be tested? It is not possible to measure directly density contrasts in the interior of the Earth, but it is possible to measure deformation of the outer surface. The ratio of geoid height to

Bradford H. Hager is Assistant Professor of Geophysics in the California Institute of Technology, Pasadena, Ca 91125.

1. Jeffreys, H. *The Earth* (Cambridge University Press, 1976).
2. Morgan, W.J. *J. Geophys. Res.* 70, 6175 (1965a); 70, 6189 (1965b).
3. McKenzie, D.P., Roberts, J.M. & Weiss, N.O. *J. Fluid Mech.* 62, 465 (1974); *Geophys. J. astr. Soc.* 48, 211 (1977).
4. Parsons, B. & Daly, S. *EOS* 60, 391 (1979).
5. Davies, G.F. *Geophys. J. astr. Soc.* 49, 459 (1977).
6. O'Connell R.J. *Tectonophysics* 38, 119 (1977).
7. Richter, F.M. & McKenzie, D.J. *Geophys. J.* 44, 441 (1978).
8. Richards, M.A. & Hager, B.H. *EOS* 62, 1077 (1981).
9. Richter, F.M. *J. geophys. Res.* 84, 6783 (1979).
10. Kaula, W.M. in *The Nature of the Solid Earth* (ed. Robertson, E.C.) 385 (McGraw-Hill, New York, 1977).
11. Chase, C.G. *Nature* 282, 464 (1979).
12. Anderson, D.L. *Nature* 297, 391 (1982).
13. McKenzie, D.P., Watts, A., Parsons, B. & Roufousse, M. *Nature* 288, 442 (1980).
14. Birch, F. *Geophys. J. R. astr. Soc.* 4, 295 (1961).
15. Nakanishi, I. & Anderson, D.L. *Bull. seismol. Soc. Am.* (in the press).

surface deformation is observable and very sensitive to the distribution of viscosity^{3,8} as well as to the depth of the convecting layer. The fact that Africa (geoid high) is on average 500 m higher than Siberia or the area around Hudson Bay (geoid lows) supports the hypothesis that the geoid in these regions results from less dense and more dense mantle respectively, with stresses transmitted efficiently to the surface of the Earth. This would be consistent with the absence of a low viscosity asthenosphere in these regions. This absence would not be surprising as these areas are ancient cratons.

McKenzie and others¹³ have looked at correlations between bathymetry, interpreted as surface deformation, and geoid anomalies on a smaller scale; these anomalies are consistent with those obtained for calculations of convection in a uniform viscosity fluid. Given the sensitivity of geoid anomalies to viscosity structure, it would be of great interest to extend the calculations to include the low viscosity asthenosphere generally supposed

to exist beneath oceanic plates to see if the geoid observations are in accord with this type of rheology.

Seismology provides additional constraints. Density and seismic velocity are, in general, positively correlated.¹⁴ At present, seismologists are making rapid progress in deciphering the three-dimensional distribution of lateral heterogeneities in seismic velocity in the Earth's interior¹⁵.

We are at the beginning of an exciting era in geodynamics. High quality geoid observations, combined with high quality determinations of surface deformation and new high resolution seismic data will provide powerful constraints which must be satisfied by the next generation of geodynamic models. Combining all three data sets will decrease the non-uniqueness inherent in each. The physics of the gravity field, which makes it effectively a null measurement, should make it possible in the near future to constrain variations in mantle rheology and the depth and scale of mantle convection. □

DNA unwinding in transcription and recombination

from L. Mark Fisher

LOCAL unwinding of the DNA helix plays an important role in many biological processes. In transcription, opening of the DNA helix by RNA polymerase allows access to the coding DNA strand and copying of the DNA message into an RNA transcript. The formation of heteroduplex DNA structures by recombination between homologous DNA molecules, catalysed by RecA (and other) proteins, must also involve unwinding of duplex DNA. Two recent papers reveal some intriguing features of DNA unwinding by *Escherichia coli* RNA polymerase¹ and RecA protein² (see this issue of *Nature*, p.185).

DNA helix unwinding may most conveniently be investigated by taking advantage of the topological properties of circular DNA. A closed circular duplex DNA is characterized by a topological parameter termed the 'linking number', which specifies the number of times the two complementary DNA strands are intertwined in the DNA circle. The linking number is an invariant and can be changed only by breaking and resealing the DNA. Therefore, if protein binding induces unwinding in one part of the helix in a closed duplex DNA circle, this must necessarily be accompanied by overwinding of the DNA elsewhere in the DNA molecule. The recent studies on unwinding by RecA protein and

by RNA polymerase have in different ways both exploited this phenomenon.

Unwinding by polymerase

RNA polymerase from *E. coli* is a pentameric complex of subunit structure $\alpha_2\beta\beta'\sigma$ and molecular weight 465,000. RNA synthesis by the enzyme proceeds via the sequential formation of four different enzyme-DNA complexes. The holo-enzyme first binds at the promoter site forming the closed promoter complex, in which the enzyme is associated with DNA sequences in their helical conformation. This complex isomerizes, unwinding and separating the DNA strands to form the open promoter (binary) complex. Exposure of the coding DNA strand allows initiation of RNA synthesis to take place. The short RNA oligomer made in the initiation complex forms an RNA-DNA hybrid with the coding strand of the DNA

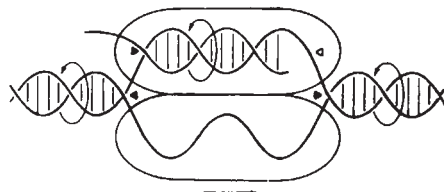
template. The sigma subunit is lost and the resulting ternary or elongation complex moves along the DNA simultaneously synthesizing the RNA transcript, which is continuously displaced from the coding DNA strand.

DNA unwinding by RNA polymerase was first unequivocally demonstrated for open promoter complexes by Wang and co-workers^{3,4}. These studies pioneered the approach of measuring DNA unwinding from the difference in linking number between DNA samples obtained by sealing a nicked circular duplex DNA with DNA ligase, in the presence and absence of the DNA binding protein⁴. Gamper and Hearst have now measured and compared DNA unwinding induced by RNA polymerase in binary, initiation and elongation complexes formed at a single prokaryotic promoter sequence in simian virus 40 DNA¹. They used gel electrophoresis to measure the change in the linking number of closed circular SV40 DNA in transcriptional complexes in which the DNA had been relaxed using a DNA nicking-closing extract. This procedure removes the overwound DNA turns arising from protein binding, thereby fixing the topological change.

The degree of DNA unwinding was found to be the same for all three transcriptional complexes and, for the elongation complex, was independent of its sequence location on DNA. The extent of DNA untwisting was deduced to be 17 ± 1 base pairs ($580^\circ \pm 30^\circ$) per polymerase molecule. The result can be compared with a value of approximately seven base pairs previously reported by Wang for open promoter complexes⁴. Although these unwinding estimates are somewhat different, the important result established by the new study is that the extent of DNA unwinding is constant throughout transcription.

Invariance of the DNA unwinding angle as the ternary complex progresses along the DNA helix suggests that the size of the unwound DNA region is tightly controlled by polymerase and is independent of the DNA sequence in the complex. A topological model that accommodates this result has been proposed to describe elongation by polymerase (Figure 1). It is suggested that RNA polymerase has two windase activities that rigidly determine the limits of the DNA unwound region and which facilitate its movement along the DNA helix. The leading unwindase opens the DNA helix which is then subsequently reformed by the lagging rewindase. These activities are coupled and result in translocation of the enzyme relative to the substrate DNA and the RNA transcript. Given the size of polymerase and that of the transcript, it seems unlikely that the ternary complex itself rotates on DNA. A more plausible and topologically feasible solution is that the DNA on either side of the unwound region rotates about the helix axis (Figure 1). Because the helix axis

Fig. 1 Topological model for the RNA polymerase elongating complex (taken from ref. 1). Filled triangles denote swivelase centres, the open triangle denotes the catalytic site for RNA synthesis. The unwound strands of the DNA are bound to polymerase, possibly to the β and β' subunits. [Figure © M.I.T. Press].



L. Mark Fisher is in the Department of Biophysics, King's College, University of London.

remains fixed, the proposed mechanism would not result in spatial revolution of the entire DNA molecule about polymerase during transcription of a circular DNA. Continuous displacement of the transcript from the DNA template is assumed to involve the rotation of the RNA-DNA hybrid. Further experimental work will be necessary to validate the details of the model and to examine its role in pausing and termination of RNA synthesis by polymerase.

Unwinding by RecA protein

The RecA protein is involved in DNA recombination and repair in *E. coli*. This 38,000 molecular weight protein has an ATP-dependent DNA strand transfer activity that promotes the formation of joint molecules between duplex DNA and either single DNA strands or double-stranded DNA with single-stranded gaps⁵. Whatever the precise mechanism of DNA strand transfer, the RecA protein must somehow unwind duplex DNA so as to catalyse the formation of heteroduplex DNA structures.

Several groups have shown that RecA protein binds cooperatively both to single-stranded and to duplex DNA⁶⁻¹¹. Complexes with duplex DNA are stabilized by the inclusion of ATP γ S a non-hydrolysable ATP analogue. These nucleoprotein complexes appear in the electron microscope as right-handed helical filaments in which the duplex DNA is elongated by about 50 per cent. This lengthening of the DNA is consistent with unwinding of the DNA helix on interaction with RecA protein¹². Stasiak and Di Capua (see p.185) have now focused on the ATP γ S-stabilized cooperative binding of RecA protein to closed circular duplex DNA². Local unwinding of this DNA by RecA protein induces a compensating overwinding (positive supercoiling) of the DNA segment not complexed with protein. The countervailing positive supercoiling of the DNA constrains the extent of protein-DNA complex formation.

The amount of RecA protein that binds to DNA depends on the negative superhelix density of the starting circular DNA. Clearly, the higher the initial negative supercoiling (underwinding) of the DNA, the more protein can bind before the binding constraint is reached. Stasiak and Di Capua used electron microscopy to

measure the fraction of the DNA circle covered by protein when negatively supercoiled DNA samples with different mean linking numbers were incubated with excess RecA protein. By extrapolation, they were able to estimate the linking number of the DNA in a closed duplex circle completely covered with RecA protein. The linking number they obtained had essentially the same value as the number of visible helical turns counted for the nicked DNA molecule fully covered with the RecA protein. These elegant studies show that the DNA helix in the complex follows the protein helix seen in the electron microscope and that RecA protein unwinds the DNA helix from the usual value of about 10.5 base pairs per turn to 18.6 base pairs per turn. Studies by Shibata's group, described in *Nature* last week (299, 86), have shown that relaxation of RecA protein — closed

circular DNA complexes by a nicking-closing extract generates highly negatively supercoiled molecules consistent with RecA mediated unwinding of the DNA.

Intercalation of protein (or conceivably ATP γ S) into the DNA helix could account for DNA unwinding in DNA-RecA protein filaments. Alternatively, the helix may be melted with complementary DNA strands separated and not base paired. Interestingly, RecA protein can itself form filamentous helical aggregates similar to those observed with DNA⁸. It seems possible, therefore, that the RecA protein filament provides a rigid framework to which the DNA is constrained to bind in an unwound conformation. The precise nature of RecA protein-DNA interactions in the filaments and the relevance of these structures in recombination await biochemical analysis. □

Murine models of multiple sclerosis

from Peter Doherty and Elizabeth Simpson

MULTIPLE SCLEROSIS is a chronic progressive debilitating disease which remains an intractable human problem. Despite many years of research there is no very clear indication as to aetiology, although a number of candidate viruses have been proposed. There is a tendency for people who develop multiple sclerosis (MS) early to have had several infections in childhood. There are also marked geographical variations, with the disease increasing in incidence with distance from the equator. Another clue is provided from HLA-typing studies which show a preponderance of HLA, A2, B7 DRW2 in MS patients from North America and northern Europe, but superimposed on this there is a Gm allotype linkage, an interesting finding in view of recent reports of Gm allotype- and HLA-linked control in autoimmune disease (Whittingham *et al. Clin. exp. Immun.* 43, 80; 1981; Uno *et al. Nature* 292, 768; 1981).

Experimental analysis has been hampered by the lack of a generally acceptable animal model. Over the past twenty years much work has been done with the acute form of experimental allergic encephalomyelitis (EAE) although the recurrent progressive pattern of MS is not reproduced. A genetic component of EAE in rats and guinea pigs is indicated, however, since the induction of disease is dramatically strain dependent. It has also been established that EAE in these species can be induced by different fragments of myelin basic protein (MBP), and that the induction of both effector and suppressor T-cell subsets can

be separated by manipulation of the immunization schedule. A further substantial advance has been the development of experimentally induced relapsing EAE, first in rats and guinea pigs, and more recently in mice, which allows more detailed genetic and immunological analysis.

Over the same period that studies on relapsing EAE have developed, other research groups have been working on virus-induced primary demyelinating processes in mice. A recent meeting* provided a timely opportunity to bring together those studying autoimmunity and virus-induced demyelination with people concerned with analysis of basic immune mechanisms and T-cell circuits in mice.

It is apparent that the relapsing EAE model in the mouse and the chronic disease process caused by some viruses such as Theiler's virus (a picornavirus) have many features which mimic the pathology of human MS. However, there are also a number of differences. The progress of relapsing EAE is not inexorable, and substantial recovery with re-myelination commonly occurs whilst it is scarcely ever seen in humans. In progressive virus models, small amounts of virus antigen persist (H. Lipton, Northwestern University Medical School, Chicago) whereas attempts to recover viruses or detect virus coding sequences that are present in MS but absent in normal human brain have never been consistently successful. Even so, the similarities are sufficient to warrant thorough analysis of cellular

Peter Doherty is in the Wistar Institute, Philadelphia, Pennsylvania 19104 and Elizabeth Simpson in the Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ.

1. Gamper, H.B. & Hearst, J.E. *Cell* 29, 81 (1982).
2. Stasiak, A. & Di Capua, E. *Nature* 299, 185 (1982).
3. Saucier, J.-M. & Wang, J.C. *Nature new Biol.* 239, 167 (1972).
4. Wang, J.C., Jacobsen, J.H. & Saucier, J.-M. *Nucleic Acids Res.* 4, 1225 (1977).
5. Radding, C.M. *Cell* 25, 3 (1981).
6. West, S.C., Cassuto, E., Muralim, J. & Howard-Flanders, P. *Proc. natn. Acad. Sci. U.S.A.* 77, 2569 (1980).
7. Stasiak, A., Di Capua, E. & Koller, Th. *J. molec. Biol.* 151, 557 (1981).
8. McEntee, K., Weinstock, G.M. & Lehman, I.R. *J. biol. Chem.* 256, 8835 (1981).
9. Di Capua, E., Engel, A., Stasiak, A. & Koller, Th. *J. molec. Biol.* 157, 87 (1982).
10. Flory, J. & Radding, C.M. *Cell* 28, 747 (1982).
11. Dunn, K., Chrysogelos, S. & Griffith, J. *Cell* 28, 757 (1982).
12. Cunningham, R.P., Shibata, T., DasGupta, C. & Radding, C.M. *Nature* 281, 191 (1979).

*The meeting was held at the Kroc Foundation, 22-26 June 1982, sponsored by the American Sclerosis Society and the Kroc Foundation and organized by Bryan Wakeman and Dale McFarlin.

immune regulation and effector function in affected animals. In fact, the main virtue of the meeting may have been in indicating that common pathways exist in the development of these apparently aetiologically unrelated models.

The influence of genetic factors is seen very clearly in both virally induced demyelinating disease and EAE mice. The SJL mouse, known to have both lymphoproliferative and autoimmune tendencies, is extremely sensitive to induction of demyelinating disease. However, a propensity to autoimmunity, itself multigenic, (A. Steinberg, NIH; A. Theofilopoulos, Scripps Foundation) is not responsible *per se*, since NZB and other autoimmune strains are not highly susceptible to EAE. The induction and expression of acute EAE in mouse is under the control of both H-2 and non-H-2 genes. Involvement of H-2 Ir genes is indicated by the effects of treating developing EAE with anti-Ia antisera (L. Steinman, Stanford University School of Medicine). Analysis of a set of recombinant inbred strains shows that one of the non-H-2 genes may be controlling a histamine receptor sensitivity to histamine release and thus susceptibility to rapid extravasation of fluid and cells in the central nervous system (CNS) of mice sensitized to MBP (D. Linthicum, USC Medical School, Los Angeles). Detailed studies of the susceptibility of relapsing EAE to a variety of mouse strains reveals that the disease may be induced in many strains other than STL but that the MBP fragment responsible varies from strain to strain (R. Fritz, Emory University). The absolute requirement for T cells is shown by the inability to induce the disease in SJL/hr⁻ mice, which lack a thymus.

The meeting also raised intriguing points about the nature of the pathological process of CNS damage itself. It is clear that the acute autoimmune disease at least is T cell mediated and seems to be caused principally by the Ly 1⁺ class of effector cells. This is not established for the virus models, but there is obviously an immune component. However, both sorts of experimental systems show instances in which the oligodendrocytes that synthesize the myelin remain intact and there is little evidence of neuronal death or axonal degeneration. Conventional immunological wisdom would have us believe that Ly 1⁺, Ia-restricted T cells are the instigators of demyelinating damage, but there is no evidence that Ia antigens are present on myelin, and previous claims that such determinants are present on the surface of oligodendrocytes do not seem to be generally accepted. Also, morphological evidence would suggest that cellular damage to the myelin sheath is a function of monocytes/macrophages rather than T cells. How then do T cells induce demyelination? Does this mean that the T-cell effect must be operating through a short-range product such as a lymphokine, lympho-

toxin or prostaglandin? As Cantor pointed out, Ly 1⁺ T cells can kill by activating macrophages to produce peroxidase, which then kills both macrophage and the Ly 1⁺ inducer T cell.

We may thus be looking at a final common pathway where persistent antigen in brain, whether virus, MBP or some other CNS component, is recognized on the surface of Ia-positive antigen-presenting cells (derived from invading monocytes) and where the demyelinating process itself is a bystander effect reflecting secreted products from the macrophage-lymphocyte interaction causing a physiological defect in oligodendrocyte function, or direct damage to the myelin sheath. If this analysis is correct, then MS could be a disease of multiple aetiology, with a common pathological process. One point that might make us doubt this possibility is that there are apparent Ir gene effects which may indicate the involvement of a single pathogen. However, this could equally reflect a common neurological sensitizing component regardless of the triggering.

What of the future? It seems that we now have reasonable murine models for investigating chronic demyelinating disease in the context of contemporary immunological and molecular biological concepts and

techniques. A clear understanding of suppressor and effector circuits in both the virus and EAE models can readily be developed. The possibility that non-MHC-restricted suppressor molecules from antigen-specific clones may be targeted onto relevant cell sets, as suggested by Cantor, can be analysed. A search can be made to see whether infection with various viruses causes development of brain-specific T and/or B cell populations using cloning and hybridoma techniques. This approach has been effective in the murine reovirus diabetes model (M. Haspel, NIH).

Do viruses trigger MS? Certainly the mild pleocytosis that is a common feature of many childhood infections should allow exposure of CNS components to immunologically competent cells which are non-specifically drawn into the CNS during mild inflammatory processes. We should, over the next 10 years, get the answer for one candidate virus, measles, since 90 per cent of children in the United States are given live vaccine at an early age and measles has virtually disappeared. In contrast, in the United Kingdom only about 15 per cent of children are vaccinated against measles. However, in the absence of any correlation between measles and MS there is ample scope for experimental analysis which has some hope of relevance.

Haemocyanins to haemerythrins

from E.F.J. van Bruggen

THE invertebrate respiratory proteins include the blue, copper-containing haemocyanins of molluscs and arthropods, the haemoglobins (or erythrocrucorins) from a variety of invertebrates, and the burgundy-coloured, non-haem iron haemerythrins of certain sipunculid worms. Although all have a common function — oxygen transport in the blood — they are typically giant molecules which show a beautiful but bewildering range of architectures. Because the molecules are so large many important findings come from electron microscopy. Unfortunately, extensive subunit heterogeneity has hitherto prevented detailed X-ray crystallography and sequencing. These problems are gradually being overcome, and a workshop was held in Leeds University* at which recent findings were discussed.

Arthropod haemocyanins have polypeptide subunits of molecular weight 70,000 which aggregate to form multimers containing up to 48 subunits. Three groups of workers reported extensive amino acid sequences of subunits of Cheliceratan

haemocyanin; those from horseshoe crab, *Tachyplesus* (T. Nemoto and T. Takagi, Tohoku University), *Limulus* (A. Riggs *et al.*, University of Texas), and scorpion, *Eurpelma* (B. Linzen *et al.*, University of Munich). All show extensive homologies, and are of particular interest in that all contain an unusual peptide, -His-His-Trp-His-Trp-His-, as part of their sequence, about 170 residues from the N-terminus. It seems probable that this forms a major part of the copper-containing active site. Nemoto's group reported 359 residues of the N-terminal sequence and about 100 residues of the C-terminal sequence, leaving about 100 residues to be sequenced.

It will be of great interest to correlate these data with the X-ray studies that are now coming to fruition. Thus, the X-ray structure of spiny lobster (*Panulirus interruptus*) haemocyanin at 4 Å resolution not only showed the tertiary structure of the binuclear copper centre, but also enabled a very precise model to be made for the quaternary structure of this hexameric molecule (W. Gaykema *et al.*, Rijksuniversiteit, Groningen). The model should

*The workshop was held in the Biochemistry Department, University of Leeds 19-22 July 1982 and sponsored by the European Molecular Biology Organization.

E.F.J. van Bruggen is in the Biochemistry Laboratory of the Rijksuniversiteit Groningen, Nijenborgh 16. 9747 AG, Groningen.

help interpret the different view of multi-hexameric molecules obtained by aperiodic averaging of electron microscopic images in combination with correspondence analysis (M. van Heel, Rijksuniversiteit, Groningen and J. Frank, New York State Department of Health, Albany). Together with information from immunoelectron microscopy using polypeptide chain-specific Fab fragments, we now have a considerably improved model for the quaternary structure of the Cheliceratan haemocyanins and their dissociation products (J. Lamy *et al.*, University of Tours; B. Linzen *et al.*, University of Munich; E. van Bruggen *et al.*, Rijksuniversiteit, Groningen).

A great deal of chemical spectroscopy now confirms previous ideas that the coppers in the active site of haemocyanin are held by imidazole nitrogens (E. Solomon, Stanford University). Interestingly, there are a number of similarities with another copper protein, *Neurospora* tyrosinase, and the enzymatic activities, known for many years to be exhibited by haemocyanins, were discussed following their re-investigation by several groups (A. Nakahara, Osaka University).

Some progress has been made with the structures of the mollusc haemocyanins and the annelid haemoglobins. For the

former, elegant work using proteases has confirmed the structure of the polypeptide in gastropod haemocyanins as containing eight linked oxygen-binding units in a polypeptide of molecular weight $> 400,000$. Cephalopod haemocyanins seem to be essentially the same (C. Gielens and R. Lontie, Katholieke Universiteit, Leuven). One of the functional units of *Helix pomatia* haemocyanin has now been crystallized. Several models for annelid haemoglobins (molecular weight up to 4×10^6) were discussed. Electron microscopy shows the molecule to be a stack of two hexagonal plates 26 nm across, but many problems remain to be resolved regarding the number of subunits, their haem content, and their arrangement in the molecule.

A study of the various developmental stages of the crustacean (N. Terwilliger, University of Oregon) gave the first indication that there may be a 'fetal haemocyanin'. This is particularly interesting as there is evidence that the variety of polypeptides which form arthropod haemocyanin (for example 12–15 in *Limulus polyphemus*) are coded by a family of mRNA molecules. The polypeptides produced appear to undergo few if any post-translational modifications (E. Wood and K. Siggins, University of Leeds). □

Oxygen in the Earth's metallic core?

from Raymond Jeanloz

THE outer core of the Earth has the properties of a molten iron alloy. Its density, however, is nearly ten per cent less than that of iron, and considerable effort has gone into determining the 'lighter component' that is alloyed with the iron. The composition of the core is of particular geochemical interest, as it provides a constraint on models of the early history of the Earth. But a major question has been: to what extent can (or must) one invoke unusual chemical properties for material existing at the 130–330 GPa (1.3–3.3 Mbar) and $\sim 4,000$ – $6,000$ K pressures and temperatures of the core? Several investigators have recently emphasized the degree to which chemical behaviour may change at such extreme conditions, as compared with our experience with the same materials at low pressures and temperatures.

At a time when sulphur was considered by many to be the most plausible candidate, Ringwood¹, reviving earlier speculations, argued forcefully that oxygen must be the light component within

the core. This might appear strange at first. After all, wüstite (ideally, $\text{Fe}^{2+}\text{O}^{2-}$) is an ionic solid, with the result that most compositions in the system Fe–FeO melt (at zero pressure) to an immiscible mixture of metallic (Fe-rich) and nonmetallic (FeO-rich) liquids. Therefore, one would hardly expect significant amounts of FeO to dissolve preferentially out of the Earth's oxide mantle and into the metallic core. In fact, experiments had already indicated that metallic iron separates out of an oxide phase at high pressures and temperatures².



where hpp indicates a high-pressure phase (structure uncertain) and ϵ is the hexagonal-closest-packed phase of iron.

We must note, however, that FeO becomes a semiconductor in the liquid state, and Ringwood's conclusion that iron and iron-oxide liquids become miscible at high temperatures ($\geq 4,000$ K) is consistent with the available data^{1,3}. That Fe in an oxide undergoes an electronic transition at high pressures has recently been verified⁴: a high-spin $\rightarrow$ low-spin transition occurs (with no change of crystal structure) at 55 GPa in Fe_2O_3 . There is no longer much

doubt that compounds within the iron–oxygen system can exhibit marked changes in chemical properties at the pressures and temperatures of the Earth's deep interior. To what extent these changes could decrease the miscibility gap between Fe and FeO liquids is unknown.

Ringwood's model for the core has been simultaneously provocative and vexing. He invokes a metallization of FeO at high pressures, such that significant amounts of the oxide can dissolve into the iron-rich core. However, the electronic-bonding properties of FeO are complex and not well understood⁵: it appears that theory (for example, quantum-mechanical band calculations) cannot help us here. On the other hand, wüstite exhibits a rich defect chemistry, with the result that stoichiometric FeO has virtually never been synthesized. Rather, the experimentalist must work with Fe_{1-x}O ($0 < x \leq 0.15$), and it is difficult to be sure that identical wüstite compositions are reproduced from one laboratory to another. Hence, not even the zero-pressure, room-temperature properties of wüstite are well documented.

The only data on Fe_{1-x}O at the conditions of the core result from shock-wave experiments⁶. Indeed, the equation-of-state measurements demonstrate that wüstite transforms from its initial state (ionically bonded, NaCl-like structure) at shock-pressures above 70 GPa. According to these data, up to 45 mol per cent FeO (~ 10 wt % O) dissolved in Fe satisfies the observed density of the core. However, neither the crystal structure nor the bonding character is determined in the shock experiments, and numerous interpretations of the transition in FeO are possible. Various authors have favoured either crystal-structured or electronic (including magnetic) changes, or both^{6–9}. A definitive answer could be provided by measurement of the electrical resistivity of wüstite at high pressures and temperatures. One group¹⁰ has found no resistivity anomalies to over 100 GPa (estimated) at

Raymond Jeanloz is in the Department of Geology and Geophysics, University of California, Berkeley, Ca 94720.

1. Ringwood, A.E. *Geochim. J.* **11**, 111 (1977).
2. Ringwood, A.E. *Origin of the Earth and Moon* (Springer, New York, 1979).
3. Mao, H.K. *Carnegie Inst. Washington Yb.* **73**, 510 (1974); Mao, H.K. & Bell, P.M. in *Energetics of Geological Processes* (eds Saxena, S.K. & Bhattacharji, S.) **236** (Springer, New York, 1977).
4. Shimoji, M. *Liquid Metals* (Academic, New York, 1977).
5. Yagi, T. & Akimoto, S. in *High-Pressure in Geophysics* (eds Akimoto, S. & Manghnani, M.H.) **81** (Center for Academic Publications, Tokyo, 1982).
6. Falicov, L.M. & Koiller, B. *J. Solid State Chem.* **12**, 349 (1975).
7. Jeanloz, R. & Ahrens, T.J. *Geophys. J.R. astr. Soc.* **62**, 505 (1980).
8. Zou, G.T., Mao, H.K., Bell, P.M. & Virgo, D. *Carnegie Inst. Washington Yb.* **79**, 374 (1980).
9. Jackson, I. & Ringwood, A.E. *Geophys. J.R. astr. Soc.* **64**, 767 (1981).
10. Navrotsky, A. & Davies, P.K. *J. Geophys. Res.* **86**, 3689 (1981).
11. Endo, S. & Ito, K. in *High-Pressure Research in Geophysics* (eds Akimoto, S. & Manghnani, M.H.) **3** (Center for Academic Publications, Tokyo, 1982).
12. Liu, L., Shen, P. & Bassett, W.A. *Geophys. J.R. astr. Soc.* **70**, 57 (1982).
13. Shen, P., Liu, L. & Bassett, W.A. in *High-Pressure Research in Geophysics* (eds Akimoto, S. & Manghnani, M.H.) **421** (Center for Academic Publications, Tokyo, 1982).

room temperature. This result does not support Ringwood's metallization hypothesis, but the effects of high temperature and of melting on the bonding character of FeO remain unknown at the pressures of the core¹¹.

Another dimension to the behaviour of iron oxides at extreme conditions has recently been emphasized by Liu and co-workers, who propose that the effects of nonstoichiometry have been underestimated^{12,13}. Extrapolating from earlier results, these workers suggest that wüstite becomes less stoichiometric with increasing pressure:



with x increasing to a value of 0.25 with pressure. Ironically, this brings us back to Mao's conclusion³ that equation (1) proceeds to the right (it is worth noting that the nonstoichiometry of wüstite is associated with microdomains of magnetite — Fe_3O_4 , or $\text{Fe}_{0.75}\text{O}$, containing both Fe^{2+} and Fe^{3+} — within the FeO structure). At face value, it would seem that this proposal works against Ringwood's hypothesis,

with Fe separating out of the oxide phase. Nevertheless, the concept that the defect density (that is, the nonstoichiometry) could increase with pressure invites further work (the existing data in support of this are not definitive).

The significance of defects lies in the control of the nonstoichiometry by oxidation–reduction conditions (that is, Fe^{2+} – Fe^{3+} equilibria). However, the oxygen fugacity can only be poorly determined in ultrahigh pressure–temperature experiments, at present, although every expectation is that this is an important variable^{2,12}. In fact, one of the major reasons for geochemical interest in the light component of the core is to better constrain the redox state of the Earth's mantle, especially during the early history of the planet. To date, much of the research in experimental geophysics has concentrated on the effects of high pressures upon mineral properties and equilibria. We are now seeing a trend emphasizing the additional influences of high temperatures and oxidation state on the chemistry of deep planetary interiors. □

Poliomyelitis — epidemiology, molecular biology and immunology

from P.D. Minor, O. Kew and G.C. Schild

At a recent World Health Organization meeting the current status of poliomyelitis in the world was discussed together with recent progress in molecular biological and immunological approaches to the study of polioviruses. Precise methods for strain characterization of polioviruses were seen as essential for epidemiological purposes and studies of vaccine safety and efficacy.

Participants from several countries (China, Japan, UK, US, The Netherlands, Sweden, Switzerland and USSR) reviewed their epidemiological experiences of poliomyelitis during the past triennium. These presentations and a general international review from data reported to WHO showed that poliomyelitis, although well controlled by the use of vaccines in developed countries, remains a major problem in the developing world^{1,2}. Epidemics of poliomyelitis in unvaccinated populations are chiefly caused by type 1 poliovirus, but in countries with well established vaccination programmes where paralytic cases have been reduced to a very small number, some of the few cases reported are associated with the use of live attenuated virus vaccines, particularly in cases involving poliovirus types 2 and 3.

Current research on the characterization of polioviruses emphasizes the biochemical analysis of poliovirus genomes, and the comparison of the antigenic characters of different strains using monoclonal antibodies.

A major advance has been provided by the determination of the complete nucleotide and deduced polypeptide sequence of a type 1 poliovirus (P1/Mahoney/USA/41)^{3,4}. The virus genome encodes a single large polypeptide which is specifically cleaved to produce the viral capsid proteins (VP1, VP2, VP3, VP4) and several non-capsid proteins, including an RNA polymerase, a viral protease⁵ and other proteins of yet unknown function. Complete processing of the polypeptide seems to occur by the combined action of cellular and virus-encoded proteases. Three distinct dipeptide sites are cleaved during processing⁵: Gln–Gly pairs are cleaved by the virus-encoded protease, Tyr–Gly pairs are probably cleaved by a cellular protease, and an Asn–Ser pair in VP0 is cleaved to produce VP2 and VP4 by an unidentified protease during the final stages of virion morphogenesis. Knowledge of the activities and specificities of the virus-encoded protease may open the way for development of a rational chemotherapy for poliovirus (and other picornavirus) infections by the use of specific inhibitors

of the processing enzymes.

Detailed knowledge of poliovirus genome sequences may also contribute to the development of improved live virus vaccines, with genetic determinants optimized for antigenicity and stabilized for attenuation. An important step in this direction is the recent determination of the total genome sequence of the Sabin type 1 attenuated vaccine strain⁶. When compared with that of the parental Mahoney strain, the genome of the attenuated derivative contained 57 base substitutions (total genome = 7,441 nucleotides), 21 resulting in amino acid changes. Twelve of the missense mutations occurred in the capsid proteins and were largely clustered in VP1, raising the possibility that attenuation in part involved modification of the virion surface. Gene sequencing is underway for other poliovirus strains, and some preliminary sequence data for Sabin type 3 were presented.

Poliovaccine development may be aided by progress in developing synthetic vaccines for another picornavirus, foot-and-mouth disease virus (FMDV). Synthetic vaccines offer the advantages of stability, increased economy of production and minimal containment hazards. Two approaches, based on the findings that the purified FMDV VP1 induces protective antibodies⁷, are being explored. The first involves expression in bacteria of a cloned VP1 gene⁸. Initial results were encouraging but the first generation of products had a much lower level of immunogenicity than conventional FMDV vaccines. The second approach involves the chemical synthesis of peptides⁹ representing sequences from VP1. Such products, when linked to a protein carrier, were more immunogenic than an equivalent mass of VP1 and stimulated neutralizing antibody and protection against virus challenge in experimental animals¹⁰.

The limited evidence available also suggests a role for VP1 in poliovirus antigenicity, although the antigenic structure of polioviruses appears to be quite complex. Data were presented indicating that isolated poliovirus polypeptides (VP1, VP2 and VP3) were each able to induce low levels of neutralizing antibody in animals and to prime a recipient for the production of neutralizing antibody on subsequent immunization with small amounts of intact poliovirus (A.L. Van Wezel and P. Van der Marel, personal communication).

Among the biochemical methods of poliovirus strain identification, ribonuclease T₁-oligonucleotide mapping of viral RNAs^{11–13} is the current technique of choice. Vaccine-related isolates can generally be readily distinguished from wild strains. Oligonucleotide maps are highly reproducible, such that good agreement exists in results and interpretations when different laboratories examine the same isolates. The technique is

P.D. Minor and G.C. Schild are in the Division of Virology, National Institute for Biological Standards, Holly Hill, Hampstead, London NW3, and O. Kew is in the Center for Disease Control, Atlanta, Georgia.

extremely sensitive in detecting minor genome variations between strains¹⁴. The tendency of polioviruses to undergo genome variation during natural infections (an excellent example of this was presented for a type 2 vaccine strain which underwent long-term replication in an immunodeficient child)¹⁵ limits the utility of this technique in detecting distant genetic relationships between strains. This same process imposes an extreme limitation on the value for strain identification of SDS-polyacrylamide gel electrophoresis of viral polypeptides. An alternative method was suggested involving restriction mapping of cDNA transcripts of viral RNA¹⁶. This approach is expected to be less sensitive than oligonucleotide mapping and consequently may reveal more distant relationships between strains which are not apparent from T₁ maps.

Several workers continue to use strain-specific absorbed sera¹⁷ for the identification of poliovirus isolates as vaccine-derived or 'wild' and this technique was considered practicable and reliable. In general, designation of isolates as vaccine-derived by this method agree well with the results of oligonucleotide mapping.

Methods of identifying strains using monoclonal antibodies were reported^{15,18-23}. International collaborative studies, organized through WHO, to examine the value of strain-specific monoclonal antibodies in strain identification and to compare the results obtained with those obtained by oligonucleotide mapping were proposed and this project is in progress. Preliminary observations from the US, UK, Netherlands and Japan suggested that the results obtained by serological and biochemical methods show excellent agreement.

Biochemical and immunological differences between types of polioviruses appear to be greatest in the portion of the genome coding for structural proteins²⁴. However, there appears to be extensive cross-reaction between types demonstrated by both immune precipitation and immunological study of separated viral proteins (VP1, VP2 and VP3) after electrophoretic transfer to nitrocellulose sheets (western blotting)²⁵. Cross-reactivity of poliovirus proteins (VP1) with sera specific for other picornaviruses (for example, echoviruses 11 and 16) were described.

The antigenic structure of polioviruses is complex. It has been known for many years that the mature infectious virus particle expresses an antigenic determinant termed D antigen not found in empty capsids or heat-inactivated virus. Similarly, empty capsids or heat-inactivated virus express antigenic determinants (C antigen) not found on infectious virus particles.

Monoclonal antibodies can react with epitopes unique to D antigen (represented by mature infectious virus) or unique to C

antigen (represented by denatured empty capsids or heated infectious virus). Other monoclonal antibodies react with particles expressing D antigen and also with particles expressing C antigen²⁶, showing that the two antigens are indeed distinct but particles expressing them share common antigenic determinants.

Only some 50 per cent of specifically D-reactive monoclonal antibodies neutralize virus infectivity, despite binding efficiently to virus²². In contrast, monoclonal antibodies reacting with both D and C particles generally neutralize virus to high titres. C-specific monoclonals do not neutralize infectivity. Attempts were made to identify the virus structural proteins reacting with the monoclonal antibodies²⁵. It was shown that only six of twenty C-reactive monoclonal antibodies against type 3 poliovirus reacted with poliovirus proteins separated by SDS-polyacrylamide gel electrophoresis and then subjected to electrophoretic transfer. Five reacted with VP1 and one with VP3. In contrast, none of 29 antibodies of D or D + C specificity, including several which were highly active in virus neutralization, reacted in this

system. The nature of the antigenic determinants of poliovirus mediating virus neutralization and protection thus remains elusive. Recent studies²⁷ have shown that monoclonal antibodies neutralizing poliovirus type 1 react with VP1 on the native virus but not after isolation from the virus particle. Determinants seem complex, being specified by the tertiary configuration of one or more polypeptides, rather than simply by amino acid sequences. However, the findings with foot and mouth disease and with isolated polypeptides from poliovirus mentioned above suggest that under certain circumstances, relatively simple structures may stimulate neutralizing antibody.

Further progress in the international conquest of poliomyelitis may depend on the development of new approaches to vaccine design leading to more cost-effective products. While modern biotechnologies are likely to have an important role in achieving this goal, and the meeting revealed a continuing high level of scientific interest, the problems are complex and require a multidisciplinary approach. □

1. Paccaud, M.F. *Wild Hlth Statistics Annual* 32, 198 (1979).
2. *Wkly epidem. Rec.* 74, 361 (1980); 23, 180 (1981); 28, 221 (1981); 42, 329 (1981); 43, 337 (1981); 48, 377 (1981), 1, 4 (1981).
3. Kitamura, N. *et al. Nature* 291, 18 (1981).
4. Racaniello, V. & Baltimore, D. *Proc. natn. Acad. Sci. U. S. A.* 78, 4887 (1981).
5. Hanecak, R. *et al. Proc. natn. Acad. Sci. U.S.A.* 79, 3973 (1982).
6. Nomoto, A. *et al. Proc. natn. Acad. Sci. U.S.A.* 79 (in the press).
7. Laporte, J. *et al. C.R. hebdo. Séanc. Acad. Sci., Paris* 276, 3399 (1973).
8. Kleid, D. *et al. Science* 214, 1125 (1981).
9. Walter, G. *et al. Proc. natn. Acad. Sci. U.S.A.* 77, 5197 (1980).
10. Bittle, J.L. *et al. Nature* 298, 30 (1982).
11. Lee, Y.F. & Wimmer, E. *Nucleic Acids Res.* 3, 1647 (1976).
12. Minor, P.D. *J. Virol.* 34, 73 (1980).
13. Nottay, B. *et al. Virology* 108, 405 (1981).
14. Aaronson, R.P. *et al. Nucleic Acids Res.* 10, 237 (1982).
15. Arita, M. *et al. Intervirology* 17 (in the press); Kew, O.M. *et al. J. gen. Virol.* 56, 337 (1981).
16. Nomoto, A. *et al. Virology* 113, 54 (1981).
17. Van Wezel, A. & Hazendonk, A. *Intervirology* 11, 2 (1979).
18. Icenogle, J. *et al. Virology* 115, 211 (1981).
19. Osterhaus, A. *et al. Intervirology* 16, 218 (1981).
20. Crainic, R. *et al. Devl Biol. Standard* 50, 229 (1982).
21. Humphrey, D. *et al. Infect. Immun.* 36, 841 (1982).
22. Ferguson, M. *et al. Lancet* (in the press).
23. Guo Ren *et al.* (in the press).
24. Romanova, L. *et al. J. gen. Virol.* 52, 279 (1981).
25. Thorpe, R. *et al. J. gen. Virol.* (in the press).
26. Schild, G.C. *et al.* (in preparation).



100 years ago

WE regret to learn of the death, at Dorpat, of Dr Kreuzwald, the publisher of old Estonian songs and poems. He was born in 1804, and studied medicine at Dorpat. When a student he began to collect songs and tales of his country-people, and in the years 1840 to 1850 he published a series of remarkable articles on Estonian antiquities, mythology, traditions, and tales. His principal work was the publication, with annotations, of the whole of the different parts of the great Estonian poem, "Kalewinoey," remarkable by its fine poetical feeling for nature and analysis of human feelings. It was translated into all the chief European languages. In 1877 Dr Kreuzwald was compelled to abandon his medical practice, and died in poverty at Dorpat.

THE *Official Messenger* of St. Petersburg announces, on September 1, that "by order of the Emperor the admission of new pupils to the course of medical training for women, at the Nicholas Military Hospital, will be discontinued after the present term. The students will be allowed to conclude their course, after which the clinical instruction for women at the hospital will be abolished." The Medical Academy for Women, the courses of which were quite equal to those of the old Universities, had 367 students. Since 1887, when the first lady students passed the examinations, 281 ladies have completed the whole course of studies.

THE English Government having sent to Egypt three of the Woolwich balloons, we may remind our readers that balloons were taken out by the French army in 1794. But it was impossible for Buonaparte to use them, the furnace for the preparation of pure hydrogen having been lost when the French fleet was annihilated by Nelson in Aboukir Bay. Conte, the engineer of the aeronauts, was created the head of Cairo arsenal, and Coutelle, their captain, was sent on a scientific mission to Upper Egypt.

from *Nature* 26, 470; September 7, 1882.

REVIEW ARTICLE

Molecular drive: a cohesive mode of species evolution

Gabriel Dover

University of Cambridge, Department of Genetics, Downing Street, Cambridge CB2 3EH, UK

It is generally accepted that mutations may become fixed in a population by natural selection and genetic drift. In the case of many families of genes and noncoding sequences, however, fixation of mutations within a population may proceed as a consequence of molecular mechanisms of turnover within the genome. These mechanisms can be both random and directional in activity. There are circumstances in which the unusual concerted pattern of fixation permits the establishment of biological novelty and species discontinuities in a manner not predicted by the classical genetics of natural selection and genetic drift.

"THE species problem is the oldest problem in biology"¹, and it is still unresolved. Despite the impressive sophistication of the mathematics of natural selection and genetic drift, there are relatively few experimental proofs of the genetic mechanisms by which species discontinuities in ontogeny and reproductive biology have arisen. There is, however, from the activities of widespread molecular mechanisms of turnover within eukaryotic genomes, a possible genetic mechanism for the origin of evolutionary novelty that is operationally different from those of natural selection and genetic drift.

Eukaryotic (and some archaebacterial²) genomes contain substantial numbers of multiple-copy families of genes and non-coding sequences, and it has been apparent for some years that many families exhibit unexpected sequence homogeneity within and between individuals of a species. A family that is shared between several species, on the other hand, often reveals substantial heterogeneity between the species.

Family homogeneity could be achieved by several molecular mechanisms which ensure that existing members of a family might be replaced in turn with a single new variant member. The rate of replacement seems to reflect both stochastic and directional changes operating within the genomes. The modes of operation of such mechanisms have unusual genetic consequences for the manner in which variation is distributed in nature and for the concerted evolutionary progress of a group of sexual individuals. Fixation of variants in a population as a consequence of stochastic and directional processes of family turnover is defined as molecular drive³.

The dynamics of natural selection and genetic drift are modelled on the premise that mutations are unitary and passive events that rely on the activities of selection and the vagaries of drift to increase in frequency in a population⁴⁻⁶. The two processes are considered to operate within mendelian populations and lead, by definition, to adaptive and non-adaptive modes of evolutionary change⁶, respectively. The widespread fixation of variants by molecular drive is different in that it is an outcome of a variety of sequence exchanges within and between chromosomes that give rise to persistent non-mendelian patterns of inheritance. Significantly, there are circumstances in which the activities of the genomic mechanisms, in spreading sequence information between chromosomes, would lead to the progressive increase of a variant through a family more or less simultaneously in each individual of a sexual population. This concerted pattern of fixation by molecular drive may provide an explanation for the origins of species discontinuities and biological novelty.

Genomic flux

The nuclear genomes of eukaryotes are subject to a continual turnover through unequal exchange, gene conversion and DNA

transposition. These processes operate both within and between chromosomes. Some families of sequences have been described aptly as fluid⁷ because of the continually changing nature of their composition, abundance and position. Continually changing families have been described in a wide range of plant and animal species^{3,8-15}. Some of the families consist of sequences with unknown function (non-genic families); others are either transcribed into functional RNA (ribosomal and transfer RNAs) or translated into proteins (for example, globins, histones, immunoglobulins, MHC proteins, actins)¹⁶⁻²⁶.

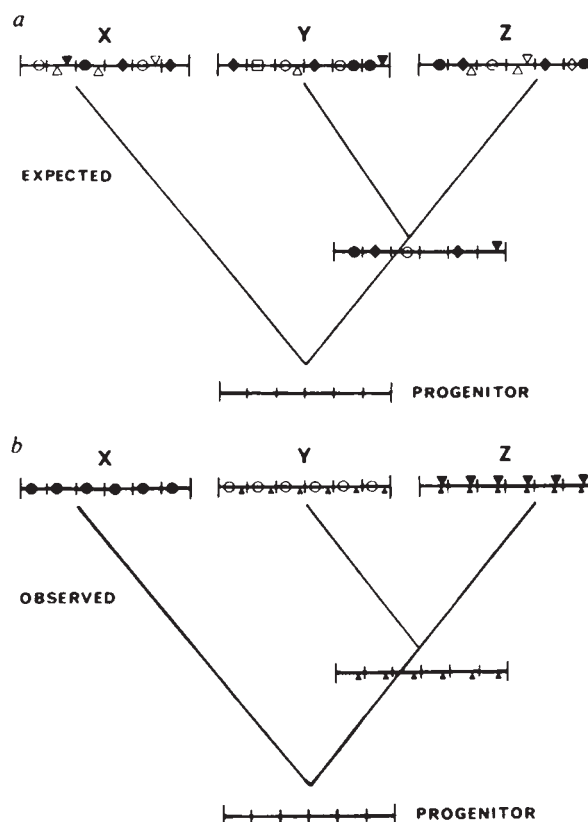


Fig. 1 Schematic representation of expected (a) and observed (b) patterns of variation in a family of sequences shared by species X, Y and Z (see also refs 29, 30). If each member is free to evolve independently, then the coefficient of identity between any two randomly chosen members within a species is expected to be the same as that between two members chosen from between the species. Observed patterns reveal a high degree of within-species homogeneity. Families may be tandem (as shown) or interspersed in the genome, and may vary from two to several million members. Families may consist of copies of genes or non-coding sequences or mixtures of both.

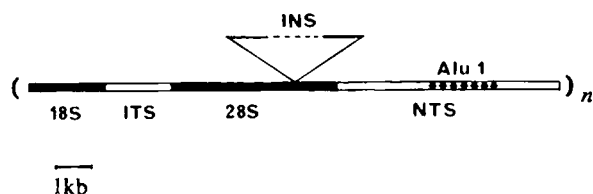


Fig. 2 A repeating unit of rDNA of *D. melanogaster*^{16,130}. 18S and 28S represent regions transcribed into functional rRNAs. INS is an insert of variable length in some repeats. ITS is an internal transcribed spacer; NTS is a non-transcribed spacer within which there are *AluI* sensitive repeats with a 240-base pair periodicity. The *AluI* motif is not present in the other six sibling species; nevertheless each species contains a diagnostic sequence variant that is present throughout the X and Y chromosome rDNA arrays (Fig. 3) (see text). n is approximately 200 in each array. There are approximately 10 *AluI* sensitive repeats in each NTS.

The molecular mechanisms of turnover help shape the patterns of variation and affect the rate of evolution of genomic families. Although there are important differences in the evolution of individual families, similarities between members of both genic and non-genic families within a species are often greater than expected if each member were evolving independently.

Figure 1 illustrates the differences between expected and observed patterns of variation in a family that is shared between three closely related species. If each member of a family is free to accumulate mutations over time then the coefficient of identity between any two members within a species would equal that of any two members chosen at random from two different species (Fig. 1a). In fact, however, many families, irrespective of their copy-number, function and genomic distribution (tandem or interspersed) display a much greater within-species than between-species homogeneity (Fig. 1b).

This characteristic pattern of variation was first described in the classic studies on interspecific differences in the sequence composition of tandem arrays of genes and spacers that comprise the ribosomal DNA (rDNA) in species of frog²⁷. The spacers reveal many mutational differences between the two species and apparently are free to diverge. Nevertheless, there are no substantial differences in sequence between the several hundred copies of the spacer-gene unit within each species. Very similar features of 'concerted' evolution have been described in other genic and non-genic families (for reviews see refs 3, 16–18, 20, 28–30). A progressive homogenization of families of sequences seems to be the fate of large proportions of species' genomes.

Molecular processes of fixation

Both stochastic and directional processes of turnover occur within nuclear genomes. The stochastic component of molecular drive results from the accidental fixation of a variant in a family and in a population through random fluctuations in the frequencies of variants within the family. Three known mechanisms can produce such fluctuations: unequal chromatid (or chromosome) exchange, gene conversion and gain and loss by transposition.

Tartof³¹ and Smith³² originally proposed that the homogenization of a family that is tandemly arrayed on a single chromosome could be a consequence of unequal chromatid exchanges.

Unequal exchange is a process which creates a sequence duplication in one chromatid (or chromosome) and a corresponding deletion in the other. There is experimental verification of mitotic and meiotic unequal exchanges in the rDNA arrays of yeast^{33,34}, and it is a possible mechanism of change in the rDNA of several species^{16,17,31,35,36}. Recurrent unequal exchange within a tandem array has the effect of expanding and contracting the copy-number of the family and the process may lead to the stochastic fixation of one or another variant member throughout the array^{30,32}.

Unequal exchange is unlikely to occur often between members of a family that are interspersed with many other sequences, particularly if they are on several nonhomologous

chromosomes. Instead it has been suggested that gene conversion is a more appropriate mechanism of homogenization within such families, and possibly also within families that are tandemly arrayed^{37–45}.

Gene conversion was first inferred from deviations from the expected 2:2 ratio of alleles in the gametes of a heterozygote individual in which the four products of an individual meiotic division remain a cluster (tetrad) so that mendelian ratios can be directly observed (tetrad analysis) (for reviews see refs 38, 46–51).

Holliday⁵² proposed that deviant ratios arise by a mismatch repair within a heteroduplex DNA region, formed during recombination between two alleles *A* and *a*, and resulting in the production of either *A + A* or *a + a*. Evidence for recombinational heteroduplexes has been obtained in *Escherichia coli*⁵³. An equal probability of conversion in each direction results in parity in the expected frequencies of *A + A* and *a + a* overall.

Gene conversion, involving up to several kilobases of DNA, has been shown by genetic and molecular analysis to occur within a chromosome^{38,50,54}, between homologous chromosomes^{45–50} and between non-homologous chromosomes^{37,55–57}. Unlike unequal exchange gene conversion does not alter the copy-number of a family; nevertheless random fluctuations in the frequencies of the direction of conversion (*Aa* → *aa* versus *Aa* → *AA*) may lead to the accidental fixation of one variant or another throughout a family. Gene conversion has been invoked as a mechanism that achieves: (1) homogeneity between pairs of relatively closely linked genes, for example pairs of mammalian globin genes^{20,41,58}, *Drosophila* heat-shock genes⁵⁹ and silk-moth chorion genes²²; (2) 'switch' in the mating type genes of yeast⁵⁴; and (3) homogeneity of amplified copies of *aprt* units in mammalian cells⁶⁰.

The directional component of molecular drive is postulated to result from persistent nonrandom exchanges of sequence information both within and between chromosomes^{3,36,61}. There are two known mechanisms that may lead to a directed homogenization of a family: one is duplicative transposition and the other is biased gene conversion.

Duplicative transposition involves the duplication and movement of a sequence from one position to another. A sequence with a high propensity to duplicate with movement of one copy to another chromosome has an increased probability of spreading in a karyotype, and hence through a population^{62–64}. Eukaryote genomes contain several such families (for reviews see refs 13, 65–67).

A bias in the direction of conversion can be observed as a departure from parity in the ratios of the expected products of conversion. The frequency of *A + A* from conversions within heterozygous cells *A + a* is consistently higher than that of *a + a*; an outcome which leads to an increasingly greater proportion of *A* in succeeding generations. The extent of bias in the conversion of different mutant sites in fungi has been observed^{39,45–47,49}. In yeast, approximately half the observed mutants depart from parity; and in some cases the departure can give rise to a 2:1 ratio in favour of the conversion of one allele by another. In other species, the deviations from parity may reveal a hundred-fold preference for the conversion of one allele by another. It has been known for several years that the conversion behaviour of a mutant is dependent on its molecular nature. Whitehouse and his colleagues^{51,68}, following from the studies of Leblon⁶⁹, have observed a predominant bias in the direction of conversion in favour of frameshift mutants containing additions of sequences, whereas single base substitutions convert with about equal frequency in each direction. Additionally, there is suggestive evidence that the 'invading' strand during heteroduplex formation is preferentially favoured during mismatch repair⁷⁰. A mutation resulting from a frameshift addition and with a high propensity to break and invade the opposing duplex might be the basis of a high conversional bias. However, the precise molecular process creating conversion disparity during recombinational mismatch repair is unknown.

Theoretical dynamics of molecular drive

Theoretical models of fixation need to consider the process of homogenizing a family of n^* copies distributed on $2n$ (diploid) number of chromosomes in a population of Ne individuals. Although the three levels of fixation are occurring simultaneously (see later), it is convenient to consider first how a variant may become fixed within an array on a single chromosome and second the fixation of the homogenized chromosomal array in the population. A thorough treatment of the dual process of fixation using the statistical dynamics of random fluctuations in unequal exchange (or gene conversion) and genetic drift has been made by Kimura and Ohta^{64,71}, and the parameters of the process have been reviewed³⁰.

It is assumed that in a randomly mating population with an effective size Ne a multiple-copy family is evolving by means of mutation, intrachromosomal unequal exchange (or unbiased gene conversion), interchromosomal recombination and genetic drift. The simplest quantity that has been studied is the 'identity coefficient' which is defined as "the probability of two randomly chosen genes belonging to the family being identical by descent"³⁰. Consideration has also been given to the expected time required to fix a variant by molecular and genetic drift (ref. 72 and T. Ohta, personal communication). The identity coefficients and fixation times can be described in terms of n^* , Ne , the rate of unequal exchange (or unbiased gene conversion), and the rate of regular recombination, together with several additional parameters concerning the mean displacement of unequal exchange and the effect of selection in stabilizing family size^{30,73} (see later).

The directional component of molecular drive, arising out of nonrandom fluctuations in the direction of conversion and duplicative transposition, has the potential to accelerate the rate of fixation^{3,36,39,49,61}. Using data on conversion frequencies and disparities between alleles of single-copy genes in fungi, Lamb and Helmi⁴⁹ have shown that, for alleles that are selectively neutral, or nearly so, fixation is largely determined by the frequency and disparity in the direction of conversion.

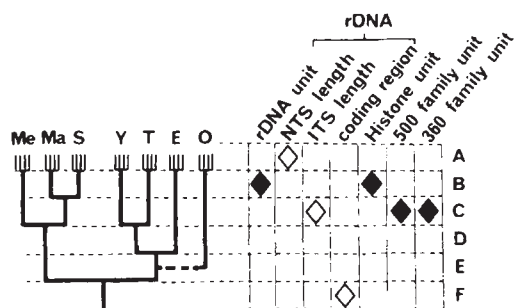


Fig. 3 Phylogeny of seven sibling species of the *melanogaster* species subgroup of *Drosophila*, based on the tempo of evolution in four repetitive families, the rDNA, the histone gene family and two non-coding families, the '500' and '360'. For details of rDNA see Fig. 2. The histone unit consists of five genes (H1, H3, H4, H2A and H2B) and four spacers (see ref. 18 for details), of which there are approximately 100 copies per haploid genome of *D. melanogaster*. The '360' and '500' families are complex in organization, large in copy-number (up to 10,000) and can be distributed over several chromosomes^{12,61}. Some regions of a family unit (or complete family unit) are placed in the phylogeny at a level (A-F) determined by the extent to which the region (or complete unit) varies between taxa. For example, the rDNA coding region is the same in all species (level F); the NTS length (see Fig. 2) of the rDNA shows intraspecific variation (level A); sequences within spacer regions of the NTS and the histone unit differ between all species but are homogeneous within each species (concerted evolution) (level B); the ITS length and various features of the '500' and '360' units vary between all species except between *D. simulans* and *D. mauritiana* (level C). $\diamond$ Represents the rate of change in sequence or length; $\blacklozenge$ represents the rate of fixation. For example, level B represents the minimal rate of fixation of the rDNA unit between the seven species due to the high rate of unequal exchange indicated by the high level of NTS length variation (level A). Although the rate of fixation of the complete rDNA unit is at level B, the rate of change in the coding regions within the unit is at F, probably as a result of selection limiting the number of mutational changes (see text and refs 3, 36, 61 for details). Me = *melanogaster*; Ma = *mauritiana*; S = *simulans*; Y = *yakuba*; T = *teissieri*; E = *erecta*; O = *orena*.

Conversion could be an important long-term factor even if the frequencies and disparities are smaller in organisms other than fungi. Unlike selection, fixation by biased gene conversion can take place without reduction in the number of offspring and, unlike genetic drift, it proceeds as easily in large as in small populations.

Nagylaki and Petes³⁹ have assessed recently the effect of biases in conversion on the sequence identity between members of a family within a single chromosome. The probability and mean time to fixation can be simulated for different values of n^* and different degrees of bias, on the assumption of one conversion event per generation between any two randomly chosen members of an array. A small conversional bias can have a dramatic effect on the probability and time for fixation. For example, the mean fixation time of a variant in an array of $n^* = 10,000$ decreases 100-fold with a conversion ratio (r) of 1.2 in favour of the variant (with parity $r = 1$; see ref. 39 for definition of r). Small biases in the direction of interchromosomal conversion are expected, similarly, to accelerate the rate for achieving homogeneity within families of large copy-number interspersed with other sequences over many chromosomes^{3,36,61}.

Many factors influence the frequency and disparity of conversion, including the nature of the interacting sequences, the genomic location and the genetic and environmental background^{45,46,49,54}. A sequence fixed as a result of a high conversional bias could not necessarily retain its favoured place indefinitely given the constantly changing spectrum of extraneous influences.

As the molecular bases of transposition, conversion and unequal exchange are not understood, it is not possible to guess what influences the DNA sequences themselves might have on their efficiency (for discussion see refs 74-76). There are hints that certain sequences (or their organization⁷⁷) might act as either hotspots or coldspots with respect to the boundaries and frequencies of conversion and unequal exchange^{22,41}. Families are expected to differ in the extent to which new variants would influence the variety of genomic mechanisms of flux at any given moment in their history. It is probable that the stochastic and directional components of molecular drive are intimately related and vary from family to family.

A comparison of expected and observed times of fixation in well-defined groups of species for families of known size and composition permits some empirical assessment of the relative contributions of the two components of molecular drive to the evolutionary process.

Assessing the tempo of fixation

To assess empirically the rates of fixation, my colleagues Enrico Coen, Tom Strachan and Steve Brown have analysed variation in multigene and non-genic families within and between species of the *melanogaster* species subgroup of *Drosophila*^{35,36,61} and in interspersed and tandem families within closely related species of rodents^{43,78,79}. Interpretations of the data viewed *in toto*³ lead to the proposition that some supplementary mechanisms are probably acting to accelerate the fixation process. Additionally, the patterns of homogeneity within families that are distributed on two or more chromosomes reveal that there is a significant exchange of sequence information between chromosomes. The existence of interchromosomal exchanges influences the extent to which individuals of a population are evolving in unison.

Coen and Strachan have examined the extent of homogeneity within and between seven sibling species of the *D. melanogaster* sub-group, for families of rDNA, histone genes and non-genic families^{35,36,61}. The data are summarized in a phylogenetic scheme (Fig. 3) based on the extent to which different regions of the families have become fixed for variants that can discriminate between the species (see ref. 3). (Details of the families are given in the legends to Figs 2 and 3.)

The families are homogeneous in each species in that they contain the same restriction enzyme sites in all members.

Different regions of the families evolve at different rates. For example, the gene regions of the rDNA (Fig. 2) are invariant (level F; Fig. 3). At the other extreme the non-transcribed spacer (NTS) of the rDNA and some coding and spacer regions of the histone repeating unit are sufficiently variant as to be different in all seven species (level B). Nevertheless, within each species the families are homogeneous for species-diagnostic sequence variants. For example, the *AluI* site present in each of the smaller repeating units within the NTS of *D. melanogaster* (Fig. 2) is present throughout all the smaller repeats within the NTSs of both the X and Y chromosomes and in many populations of this species distributed worldwide. It is not present, however, in the NTSs of the other species. The rate of fixation in the rDNA and histone families is placed accordingly at level B (see refs 3, 36 for further explanation).

High variability in the lengths of NTSs and in the copy-number of each length within individual rDNA arrays³⁵ (level A) indicates that frequent unequal exchanges probably play a significant role in the fixation of variant NTSs.

Using the statistical dynamics of family and population size fluctuations, Ohta (ref. 72 and personal communication) and my colleagues and I^{3,36} have evaluated the number of generations required to fix a variant in the rDNA family where $n^* \approx 200$ per chromosome, $N_e = 10^5$, there is an upper limit in the rate of unequal exchange of 10^{-3} – 10^{-4} per generation³⁵ and an unequal displacement of 20%⁷² and 10%³⁶ of the array length. Both estimates arrive at a mean time for fixation of the order of 10^7 generations. The fixation time is close to estimates of a time of separation of 10^6 – 10^7 generations between *D. simulans* and *D. melanogaster*^{59,80} which are based on genetic distances (Nei's measure, *D*, ref. 81). The correlation between *D* and real time is extrapolated from the observed relationship between genetic distances of Hawaiian species of *Drosophila* and the age of the volcanic islands⁸². The fixation time estimates are probably an underestimation in that the rDNA is distributed on the X and Y chromosomes, both of which need to be maintained in equal frequencies during genetic drift. Additional retardation in the rate of fixation would result if the mean unequal displacement is similar to the 5% displacement observed in the rDNA of yeast³³. Although assessments of fixation times are accompanied by large variances, it is possible that both directional and stochastic processes have been involved in the fixation of rDNA variants in the relatively shorter time separating the appearance of *D. simulans* and *D. mauritiana*³ (see Fig. 3). This possibility is open to experimental investigation.

Some directional process may also be required to fix variants in the '360' and '500' families having up to 10,000 members which can be distributed over several chromosomes⁶¹. These families have been fixed for diagnostic variants in all species except *D. simulans* and *D. mauritiana* (level C). If such non-genic families can tolerate large fluctuations in copy-number by unequal exchange then the time taken for fixation by stochastic processes is correspondingly reduced^{30,73}. Wide polymorphic fluctuations in copy-number in these families (compare level A) are as yet unknown, but this possibility can be tested experimentally. The distributions of families to many chromosomes might pose additional problems with the effect of slowing down the rates of homogenization by drift^{3,61}.

A family of $n^* = 20,000$ members that is interspersed with other sequences on up to 40 chromosomes and does not substantially alter in copy-number (the MIF-1 family)^{43,83,84} has revealed also the fixation of diagnostic variants between species of *Mus*⁴³. Palaeontological⁸⁵ and genetic distance⁸⁶ estimates put the time since separation of some of the species at $1\text{--}2 \times 10^6$ yr. Simulated estimates are available of the number of conversion generations within a single chromosome lineage that are needed to homogenize a family of $n^* = 20,000$ through stochastic fluctuations in unbiased intrachromosomal conversions, assuming one conversion per generation³⁹. For example, the mean time to fixation when $n^* = 20,000$ is approximately 4.00×10^8 conversion generations, or 10^8 yr, assuming four

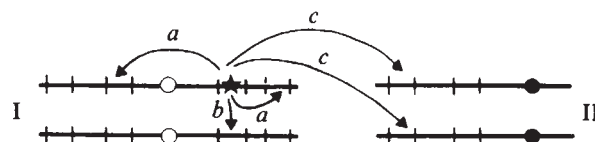


Fig. 4 Schematic representation of pathways of transfer of sequence information of a variant member. *a*, Intrachromosomal homogenization either by unequal exchange or biased/unbiased gene conversion. *b*, *c*, Pathways of homogenization between homologous (I) or nonhomologous (II) chromosomes. Non-reciprocal sequence transfer between chromosomes could result from either duplicative transposition or biased/unbiased gene conversion. The relative rates of transfer between *a* and *b* + *c* and the rate of chromosome turnover (generation time) determine the extent to which a simultaneous change in the ratio of new to old variants occurs in each individual of a population (see text).

generations per year in the rodent species. The model is based on an individual chromosome lineage and does not consider the requirement to fix a variant on 40 chromosomes in a population of N_e individuals³⁹. Such considerations would increase the number of conversion generations required for fixation. However, given this minimum calculation, to fix a variant in a family by stochastic processes within 10^6 yr, 100 conversions per individual per generation would be required. A small and persistent conversional advantage of a variant, for example a conversional ratio (*r*) of 1.2 (see ref. 39 for definition), would reduce the number of required conversions per individual to one per generation, in order to fix within 10^6 yr. Such estimates clearly are open to empirical investigation.

It is reasonable to postulate, for the tandem and dispersed genic and non-genic families being considered in *Drosophila* and *Mus*, that known mechanisms of directional change may have made some contribution to the distinctive patterns of species homogeneity and may have accelerated, somewhat, the process of fixation along the lines modelled by Lamb⁴⁹ and Petes³⁹ and their colleagues. Biased conversions could be taking place within tandem arrays as well as between dispersed members. Both stochastic and directional processes could be responsible for the fixation of variants in such families in these groups of species. The relative contribution of either process has yet to be assessed. It is possible that the unexpected high levels of homogeneity of many genic and non-genic families in other species is also an outcome of a combination of stochastic and directional mechanisms of turnover within nuclear genomes.

The limited independence of chromosome evolution

The analysis of sequence identity between members of families that are distributed over several chromosomes reveals that there are considerable levels of exchange of sequence information between chromosomes in many species. For example, the occurrence of the *AluI* sensitive repeats throughout the rDNA units (Fig. 2) on both the X and Y chromosomes of *D. melanogaster* must be due to exchanges between these two sex chromosomes^{35,36}. There is molecular and cytological evidence of exchanges between the chromosomes (E. Coen and G. Dover, in preparation). Because of such exchanges the evolution of the individual arrays of rDNA are not evolving entirely independently^{3,36}.

Restrictions on the independence of chromosome evolution are also revealed by the presence of smaller subfamilies of members within large families, identified by restriction sites not present in the rest of the family^{14,78,79,87–90}. Subfamilies are generally non-overlapping and are indicative of a phenomenon in which different segments of a large family are being homogenized for different variants. Surprisingly, individual subfamilies in both an abundant tandemly-arrayed family and the interspersed MIF-1 family in rodents have been found not to be localized to a single chromosome. Most subfamilies of the two families seem to be present on several chromosomes in the same proportion (refs 43, 78, and S. Brown, unpublished results). Similarly, recent studies on the chromosomal distribution of variant subfamilies in the rDNA family of humans reveal

the presence of each subfamily on all five of the chromosomes containing rDNA⁹¹.

Although there are clear examples and valid arguments for the independent evolution of either subfamilies or individual members of families that have become separated on either the same or nonhomologous chromosomes^{23,92} (for example, with respect to the organizational and functionally distinct arrays of insect chorion genes^{22,93}, sea-urchin histone genes^{18,19,94}, and *Xenopus* 5S rRNA genes^{17,95,96}), there are nevertheless many karyotypically dispersed genic and nongenic families in which interchromosomal exchanges of sequences must be taking place.

Sequence homogeneity between chromosomes effected by stochastic and directional processes of unequal exchange, gene conversion and duplicative transposition, has specific and unusual genetic consequences.

Unusual genetic consequences of molecular drive

To understand the unusual genetic consequences of homogenizing family members between homologous and nonhomologous chromosomes, it is necessary to appreciate changes in the frequencies of a mutation in three hierarchical sets of populations.

At the lowest level is the population of sequences, often dispersed on two or more chromosomes.

The second level is represented by the population of chromosomes shared between a group of sexually reproducing individuals, that is, $2n \times N_e$. In a sexual population, with a haploid/diploid alternation of generations, the chromosomes can be considered as belonging to one pool, or population, of chromosomes as a consequence of random chromosome assortment during meiosis and random fusion of gametes. How far the turnover of chromosomes generates a random population of chromosomes (that is, the frequency at which a chromosome may become associated with any other in the same individual), depends on the reproductive biology and generation time of the species.

The third level is the traditionally understood population of the individual diploid organisms.

Molecular drive involves changes in frequency of a variant member of a family at all three population levels. It is a process which could lead to a simultaneous increase of a variant in all individuals as a consequence of the large disparity between the rates of turnover of sequences and chromosomes.

The differences in rates between intrachromosome (a) and interchromosome ($b+c$) homogenization and the generation time affect the overall rate and pattern of fixation of a variant in a population of individuals (Fig. 4). It is unlikely, given the experimental assessments of the intra- and interchromosomal rates of unequal exchange and gene conversion^{33,35,36,38,49,50,57,61}, that a family would become fixed for a variant within the lifetime of an individual. The coefficient of identity between any two chromosomes taken at random from within a population, with respect to the numbers of variant repeats they have acquired, would depend on the rates of intra- and interchromosomal homogenization in relation to the rates at which chromosomes within the population are being effectively randomized at each generation. The slow rate of homogenization between chromosomes relative to the fast rate of chromosomes turnover within a population (generation time) could ensure that, at any given time, there is in each individual the same average ratio of new to old variants for a particular family^{13,97}. Hence during the period of transition the differences between individuals would be kept to a minimum with respect to the changing sequence composition of their chromosomes. Accordingly, processes of sequence homogenization between chromosomes would increase the population mean of the ratio of new to old variants (and possibly the phenotype, see below), but not the variance. Different families would experience different transition times depending on the factors affecting the rates of intra- and interchromosomal mechanisms of

homogenization and the generation times of the organisms. The dynamics of change at the three interrelated levels of population (sequences, chromosomes and individuals) are complex and a mathematical description is in preparation⁹⁸.

Embodied in the concept of molecular drive is the idea of "selfishness" as previously defined^{62,63} and subsequently refined^{74,75}. Molecular drive is a more inclusive notion in that it is seen as an outcome of stochastic and directional mechanisms of turnover in the genome that aid in the fixation of all manner of sequences irrespective of their copy-number, dispersion patterns or function. The mechanism suggested as the basis for the accumulation of selfish DNA is that of duplicative transposition^{62,63}, which was originally postulated to account for the increase in copy-number of transposed (mobile) elements. Studies of mobile elements in *Drosophila* and yeast show that the copy-number of mobile families are relatively small and do not fluctuate widely^{7,13}. They are small enough to be potentially eliminated by stochastic processes^{92,100}.

Biological consequences of molecular drive

Several important considerations arise from genomic mechanisms of interchromosomal sequence exchange having the potential to impart genetic cohesiveness to a large panmictic population, both during and after a period of change^{3,97}.

Let us consider a family of genes or non-genic sequences affecting ontogenetic growth patterns of an organism. If the rate of intrachromosomal homogenization is high then the ensuing heterozygosity between the new and old chromosomal arrays might reduce the fitness of an individual organism, leading to the loss of the array. In this circumstance the evolutionary progress of the array, as a single unit, can be monitored by classical mendelian genetics and darwinian selection. If, however, interchromosomal family homogenization is taking place, but at a rate much slower than the rate at which the chromosomes effectively constitute a random population (short generation time), then there would be a concerted change in the ontogenetic growth patterns of a group of individuals. For example, the individuals might be turning into monsters, yet there would be no significant increase in variance in the population in the degree of monstrosity. The fixation of variant sequences on different chromosomes by stochastic and directional mechanisms would induce a concerted phenotypic change in a group of individuals without the concomitant effect of disturbing relative differences in individual fitness. This is based on the assumption that small differences between individuals in the ratio of new to old variants of a multigene family during fixation have small effects on relative fitnesses. If a progressive concerted change in phenotype were ultimately to affect the viability of the organisms, then all the individuals would be affected as a group. It is possible that some aspects of the phenotypic cohesiveness of present-day species have developed along such lines (for example, see ref. 101; A. Garcia-Bellido, personal communication).

The biological consequences of molecular drive may take effect, however, at a higher level. Molecular drive may contribute to a mode of accidental speciation^{3,97,102} for maximizing the genomic and phenotypic cohesiveness within a population maximizes biological discontinuities between populations. Whether a genomic dissimilarity between any two individuals from two such populations becomes a major problem of reproductive or developmental incompatibility (speciation), depends on the nature of the sequences undergoing concerted homogenization in each population. Clearly, the widespread fixation of some variants in multigene families such as histones, rDNA, globins, immunoglobins and so on, would affect the phenotype. If, in addition, future studies confirm that many non-genic families affect DNA transcription and transcript processing^{8,103,104}, or chromatin structure and chromosome behaviour^{9,12,15,97,105-112}, then the dual process of generating intrapopulation cohesion and interpopulation discontinuities would be of considerable evolutionary significance. Hitherto, all theoretical exegeses on the genetics of species discontinuities

have considered the effects of natural selection and genetic drift operating within mendelian populations^{4-6,113-122}. A progressive concerted change in phenotype of a group of individuals effected by internal processes of genomic turnover between chromosomes could not have been foreseen in the light of classical considerations of the dynamics of population change based on individual variation and "passive" allelic mutations.

The phenomenon of "hybrid dysgenesis" between populations of *D. melanogaster* in which differential accumulation of some mobile elements has occurred^{13,100,123-125} is an example of the fixation of sequences by drive (in this case by duplicative transposition between chromosomes) leading to genomic dissimilarities and subsequent incompatibilities. Additionally, it has been known for some time that hybrids between *D. simulans* and *D. melanogaster* exhibit sterility and phenotypic disharmonies as a result of numerous small genetic differences between the chromosomes of their respective karyotypes^{126,127}. "Hybrid sterility is but one expression of this cryptic divergence, which need not in itself have had a selective value"¹²⁶.

It could be that the common ontogeny of individuals of a species is due to the unfolding effects of accidentally acquired sets of variant families and is of limited adaptive significance. The clustering of species to a remarkably few of all possible modes of development (a point expressed succinctly by Lewontin¹²⁸) might reflect in part the slow rates of production and fixation of family variants having significant ontogenetic effects. A cohesive change in species phenotype by molecular drive might be rare relative to the constancy of phenotype: an interesting parallel with 'punctuated equilibria'¹²⁹. Family turnover

is continuous and is analogous to the turnover of new for old banknotes: the introduction of a new design is, however, a rare event.

Testing the theory of the origins of evolutionary novelty and species formation by means of molecular drive requires an intense study of the modes and rates of operation of the candidate genomic mechanisms. The important differential in the rates of intra- and interchromosomal sequence homogenization relative to the rate of chromosome turnover and the relative contributions of stochastic and directional components of molecular drive needs to be analysed in depth. The numbers of affected families (genic and non-genic), their chromosomal distribution and their phenotypic effects need to be assessed in genetically well-defined groups of species and populations.

Molecular drive is not an alternative to the evolutionary processes of natural selection and genetic drift in the fixation of mutational variants of single-copy genes (although single genes are not excluded from the activities of directed conversion and may also be driven⁴⁹); it constitutes a third mode of evolution that would of necessity be subject to natural selection and genetic drift in an interesting variety of real biological situations. There is no simple solution to the species problem.

I thank Richard B. Flavell, George Miklos, John Thoday, Leslie Orgel, Richard Lewontin, Francis Crick, Michael J. D. White, Claudio Schazzocchio, Alex Jeffreys, Motoo Kimura, Tomoko Ohta, Walter Bodmer and Steve Gould for stimulating conversations, and most importantly my colleagues Enrico Coen, Steve Brown and Tom Strachan for discussing their data and their ideas on genomic turnover.

1. Dobzhansky, T. in *Molecular Evolution* (ed. Ayala, F. J.) 95-105 (Sinauer, Boston, 1976).
2. Sapienza, C. & Doolittle, W. F. *Nature* **295**, 384-389 (1982).
3. Dover, G. A. et al. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 343-374 (Academic, London, 1982).
4. Fisher, R. A. *The Genetical Theory of Natural Selection* (Dover, New York, 1930).
5. Wright, S. *Proc. 6th Int. Congr. Genet.* **1**, 356-366 (1932).
6. Kimura, M. *Scient. Am.* **251**(5), 94-104 (1979).
7. Young, M. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6274-6278 (1979).
8. Davidson, E. H. & Britten, R. J. *Science* **204**, 1052-1059 (1979).
9. Flavell, R. B. *Rev. Pl. Physiol.* **31**, 569-596 (1980).
10. Flavell, R. B. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 301-324 (Academic, London, 1982).
11. Singer, M. F. *Int. Rev. Cytol.* (in the press).
12. Brutlag, D. L. *Rev. Genet.* **14**, 121-144 (1981).
13. Spradling, A. C. & Rubin, G. M. *Rev. Genet.* **15**, 219-264 (1981).
14. Dover, G. A., Strachan, T. & Brown, S. D. M. in *Evolution Today* (eds Scudder, G. G. E. & Reveal, J. L.) 337-349 (Hunt Institute, Pittsburgh, 1981).
15. Miklos, G. L. G. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 41-68 (Academic, London, 1982).
16. Long, E. H. & David, I. B. *Rev. Biochem.* **49**, 727-764 (1980).
17. Fedoroff, N. *Cell* **16**, 697-710 (1979).
18. Kedes, L. H. A. *Rev. Biochem.* **48**, 837-870 (1980).
19. Hentschel, C. C. & Birnstiel, M. L. *Cell* **25**, 301-313 (1981).
20. Jeffreys, A. J. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 157-176 (Academic, London, 1982).
21. Davidson, E. H., Thomas, T. L., Scheller, R. H. & Britten, R. J. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 177-192 (Academic, London, 1982).
22. Jones, W. C. & Kafatos, F. C. *Cell* **22**, 855-867 (1980).
23. Bodmer, W. F. *Am. J. hum. Genet.* **33**, 664-682 (1982).
24. Maniatis, T., Fritsch, E., Lauer, J. & Lawn, R. M. *Rev. Genet.* **14**, 145-178 (1980).
25. Rabbitts, T. H., Bentley, D. L., Forster, A., Milstien, S. G. & Matthysens, G. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 205-218 (Academic, London, 1982).
26. Zachau, H. G., Höcht, J., Neumaier, P. S., Pech, M. & Schnell, H. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 193-204 (Academic, London, 1982).
27. Brown, D. D., Wensink, P. C. & Jordan, E. J. *molec. Biol.* **63**, 57-73 (1972).
28. Dover, G. A. & Coen, E. S. *Nature* **290**, 731-732 (1981).
29. Hood, L., Campbell, J. H. & Elgin, S. C. R. *Rev. Genet.* **9**, 305-353 (1975).
30. Ohta, T. *Evolution and Variation in Multigene Families* (Springer, Berlin, 1981).
31. Tartof, K. *Cold Spring Harb. Symp. quant. Biol.* **38**, 491-500 (1974).
32. Smith, G. P. *Cold Spring Harb. Symp. quant. Biol.* **38**, 507-513 (1974).
33. Szostak, J. W. & Wu, R. *Nature* **284**, 426-430 (1980).
34. Petes, T. D. *Cell* **19**, 765-774 (1980).
35. Coen, E. S., Thoday, J. M. & Dover, G. A. *Nature* **295**, 564-568 (1982).
36. Coen, E. S., Strachan, T. & Dover, G. A. *J. molec. Biol.* **158**, 17-35 (1982).
37. Scherer, S. & Davis, R. W. *Science* **209**, 1380-1384 (1980).
38. Jackson, J. A. & Fink, G. *Nature* **292**, 306-307 (1981).
39. Nagylaki, T. & Petes, T. D. *Genetics* **100**, 315-337 (1982).
40. Baltimore, D. *Cell* **24**, 592-594 (1981).
41. Shen, S. H., Slightom, J. L. & Smithies, O. *Cell* **26**, 191-203 (1981).
42. Birky, C. J. & Skavaril, R. V. *Genet. Res.* **27**, 249-265 (1976).
43. Brown, S. D. M. & Dover, G. A. *J. molec. Biol.* **150**, 441-446 (1981).
44. Ohta, T. *Genet. Res.* **30**, 89-91 (1977).
45. Edelman, G. M. & Gally, J. A. in *Neurosciences: Second Study Program* (ed. Schmitt, F. O.) 962-972 (Rockefeller University Press, New York, 1970).
46. Fincham, J. R. S., Day, P. R. & Radford, A. *Fungal Genetics* 4th edn (Blackwell, Oxford, 1979).
47. Fogel, S. R., Mortimer, K., Lusnak, K. & Tavares, F. *Cold Spring Harb. Symp. quant. Biol.* **43**, 1325-1341 (1978).
48. Finnerty, V. in *The Genetics and Biology of Drosophila* (eds Ashburner, M. & Novitski, E.) Vol. 1a, 331-347 (Academic, London, 1976).
49. Lamb, B. C. & Helms, S. *Genet. Res.* **39**, 199-217 (1982).
50. Klein, H. L. & Petes, T. *Nature* **289**, 144-148 (1981).
51. Whitehouse, H. L. K. *Genetic Recombination—Understanding the Mechanisms* (Wiley, New York, in the press).
52. Holliday, R. *Genet. Res.* **5**, 282-304 (1964).
53. Dasgupta, C., Wu, A. M., Kahn, R., Cunningham, R. P. & Radding, C. M. *Cell* **25**, 507-516 (1981).
54. Klar, A. J. S., Stratern, J. N. & Hicks, J. B. *Cell* **25**, 517-524 (1981).
55. Ernst, J. F., Stewart, J. W. & Sherman, F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6334-6338 (1981).
56. Munz, P. & Leopold, U. in *Molecular Genetics of Yeast* (eds Von Wettstein, D. et al.) 264-275 (Munksgaard, Copenhagen, 1981).
57. Mikus, M. D. & Petes, T. D. *Genetics* (in the press).
58. Leibhaber, S. A., Goossens, M. & Kan, Y. W. *Nature* **290**, 26-29 (1981).
59. Brown, A. J. L. & Ish-Horowicz, D. *Nature* **290**, 677-682 (1981).
60. Roberts, J. M. & Axel, R. *Cell* **29**, 109-119 (1982).
61. Strachan, T., Coen, E. S., Webb, D. A. & Dover, G. A. *J. molec. Biol.* **158**, 37-54 (1982).
62. Orgel, L. E. & Crick, F. H. C. *Nature* **284**, 604-607 (1980).
63. Doolittle, W. F. & Sapienza, C. *Nature* **284**, 601-603 (1980).
64. Kimura, M. & Ohta, T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4001-4005 (1979).
65. Moveable Genetic Elements *Cold Spring Harb. Symp. quant. Biol.* **45** (1981).
66. Finnegan, D. J., Will, B. M., Bayev, A. A., Bowcock, A. M. & Brown, L. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 29-40 (Academic, London, 1982).
67. Young, M. W. & Schwartz, H. E. *Cold Spring Harb. Symp. quant. Biol.* **45**, 629-640 (1981).
68. Yu-Sun, C., Wickramaratne, M. R. T. & Whitehouse, H. L. K. *Genet. Res.* **29**, 65-81 (1977).
69. Leblond, G. *Molec. gen. Genet.* **115**, 36-48 (1972).
70. Markham, P. & Whitehouse, H. L. K. *Nature* **295**, 421-423 (1982).
71. Ohta, T. *Genet. Res.* **31**, 13-28 (1978).
72. Ohta, T. *Genet. Res.* (submitted).
73. Ohta, T. *Genetics* **99**, 555-571 (1982).
74. Orgel, L. E., Crick, F. H. C. & Sapienza, C. *Nature* **288**, 645-646 (1980).
75. Dover, G. A. & Doolittle, W. F. *Nature* **288**, 646-647 (1980).
76. Dover, G. A. *Nature* **285**, 618-620 (1980).
77. Hentschel, C. C. *Nature* **295**, 714-716 (1982).
78. Brown, S. D. M. & Dover, G. A. *Nucleic Acids Res.* **8**, 781-792 (1980).
79. Brown, S. D. M. & Dover, G. A. *Nature* **285**, 47-49 (1980).
80. Eisses, K. T., Van Dijk, H. & Van Delden, W. *Evolution* **33**, 1063-1068 (1979).
81. Nei, M. *Molecular Population Genetics and Evolution* (North-Holland, Amsterdam, 1975).
82. Carson, H. L. *Nature* **259**, 395-396 (1976).
83. Maniatis, T. *Nucleic Acids Res.* **8**, 3247-3259 (1980).
84. Heller, R. & Arnheim, N. *Nucleic Acids Res.* **8**, 5031-5042 (1980).
85. Brothwell, D. in *Biology of the House Mouse* (ed. Berry, R. J.) 1-11 (Academic, London, 1981).
86. Thaler, L., Bonhomme, F. & Britton-Davidian, J. in *Biology of the House Mouse* (ed. Berry, R. J.) 27-39 (Academic, London, 1981).
87. Scheller, R. H. et al. *J. molec. Biol.* **149**, 15-39 (1981).
88. Horz, W. & Zachau, G. G. *Eur. J. Biochem.* **45**, 501-510 (1977).
89. Dennis, E. S., Dunsmuir, R. & Peacock, W. J. *Chromosoma* **79**, 179-198 (1980).
90. Southern, E. M. *J. molec. Biol.* **94**, 51-69 (1975).
91. Krystal, M., D'Ustachio, P., Ruddle, F. H. & Arnheim, N. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5744-5748 (1981).
92. Selker, E. U. et al. *Cell* **24**, 819-828 (1981).
93. Jones, W. C. & Kafatos, F. C. *J. molec. Biol.* (in the press).
94. Simpson, R. T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6803-6807 (1981).
95. Ford, P. J. & Brown, R. D. *Cell* **8**, 485-493 (1976).
96. Ford, P. J. & Southern, E. M. *Nature new Biol.* **241**, 7-12 (1973).
97. Dover, G. A. in *Mechanisms of Speciation* (eds Barigozzi, C., Montalenti, G. & White, M. J. D.) (Liss, New York, in the press).
98. O'Donald, P., Friday, A. & Dover, G. A. (in preparation).
99. Fink, G., Farabaugh, P., Roeder, G. & Chaleff, D. *Cold Spring Harb. Symp. quant. Biol.* **45**, 575-580 (1981).

100. Engels, W. R. *Cold Spring Harb. Symp. quant. Biol.* **45**, 561–565 (1981).
101. Garcia-Bellido, A. *Symp. Soc. dev. Biol.* (in the press).
102. Dover, G. A. *Nature* **272**, 123–124 (1978).
103. Jelinek, W. R. *et al. Proc. nat. Acad. Sci. U.S.A.* **77**, 1398–1402 (1980).
104. Davidson, E. H. & Posakony, J. W. *Nature* **297**, 633–635 (1982).
105. Bennett, M. D. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 239–262 (Academic, London, 1982).
106. Macgregor, H. C. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 325–342 (Academic, London, 1982).
107. Hotta, Y. & Stern, H. *Cell* **27**, 309–320 (1981).
108. Manuelidis, L. in *Genome Evolution* (eds Dover, G. A. & Flavell, R. B.) 263–286 (Academic, London, 1982).
109. Smith, G. R. *Cell* **24**, 599–600 (1981).
110. Wittig, S. & Wittig, B. *Nature* **297**, 31–38 (1982).
111. Weisbrod, S. *Nature* **297**, 289–295 (1982).
112. Darnell, J. E. *Nature* **297**, 365–371 (1982).
113. White, M. J. D. *Modes of Speciation* (Freeman, San Francisco, 1978).
114. Dobzhansky, T. *Genetics of the Evolutionary Process* (Columbia University Press, New York, 1970).
115. Lande, R. *Am. Nat.* **116**, 463–479 (1980).
116. Templeton, A. R. *A. Rev. ecol. Syst.* **12**, 23–48 (1981).
117. Thoday, J. M. *Proc. R. Soc. B* **182**, 109–143 (1972).
118. Maynard-Smith, J. *Am. Nat.* **100**, 637–650 (1966).
119. Clarke, B. *Am. Nat.* **100**, 389–402 (1966).
120. Stebbins, G. L. & Ayala, F. J. *Science* **213**, 967–971 (1981).
121. Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York, 1974).
122. Mayr, E. *Animal Species and Evolution* (Harvard University Press, Cambridge, Massachusetts, 1963).
123. Bregliano, J. C. & Kidwell, M. C. in *Mobile Genetic Elements* (ed. Shapiro, J. A.) (Academic, New York, in the press).
124. Rubin, G. M., Kidwell, M. G. & Bingham, P. M. *Cell* **29**, 987–994 (1982).
125. Bingham, P. M., Kidwell, M. C. & Rubin, G. M. *Cell* **29**, 995–1004 (1982).
126. Muller, H. J. & Pontecorvo, G. *Genetics* **27**, 157 (1941).
127. Muller, H. J. & Pontecorvo, G. *Nature* **146**, 199–200 (1940).
128. Lewontin, R. C. *Am. Nat.* (in the press).
129. Gould, S. J. & Eldridge, N. *Palaeobiology* **3**, 23–40 (1977).
130. Glover, D. M. *Cell* **26**, 297–298 (1981).

ARTICLES

A satellite altimetric geoid in the Philippine Sea

Ki-iti Horai

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

A detailed geoidal map of the Philippine Sea, constructed from GEOS 3 altimeter data, revealed that the average geoidal height in the basins of the Philippine Sea decreases as the age of the basins increases. The geoidal height–age relationship is comparable with that of mid-oceanic ridges, suggesting that the cooling of the marginal sea basin's lithosphere after its formation is similar to that of the oceanic lithosphere created at the mid-oceanic ridges. The data appear to be consistent with the back-arc opening model. However, the corner flow model is also compatible with the smoothly varying geoid over the entire Philippine Sea.

THE advent of satellite altimetry has allowed a detailed delineation of sea-surface height in oceanic areas. In sea areas where oceanographic disturbances do not predominate, the height of the sea surface measured from the Earth's standard ellipsoid provides an excellent approximation of a marine geoid. As a first step in studying the mechanical and thermal characteristics of the marginal basins of the West Pacific, a geoid map of the Philippine Sea was constructed using GEOS 3 altimeter data. The Philippine Sea was chosen because it developed behind the Mariana Island Arc that is, "in many respects, the simplest and easiest of all island arc systems to study"¹.

Data adjustment

The altimeter data consist of a set of sea-surface height values measured along the satellite's track. At cross-overs between northbound (ascending) and southbound (descending) tracks, sea-surface height values are expected to coincide. In reality, however, they disagree owing to the satellite's orbit tracking error, the altimeter's instrumental bias and the noise due to variable sea-surface condition. A consistent field of sea-surface

height is obtained by making corrections for the track data's erroneous bias and tilt, so that the r.m.s. residual from all intersections in the area becomes minimal². In the area of study covering the Philippine Sea, there are 1,575 intersections among 48 ascending and 58 descending tracks. The r.m.s. cross-over residual was more than 15 m for the original uncorrected data. That was decreased to 1.5 m by removing the data's erroneous bias by a method of iterative correction³. Further tilt correction brought the r.m.s. residual to 46 cm. A geoidal map with a contour interval of 1 m was prepared based on this final corrected data (Fig. 1a).

Figure 1b is a map of the 'residual'⁴ geoid that was derived from the total geoid, as shown in Fig. 1a, by removing the spherical harmonic component of GEM-10 geoid⁵ expanded to the order of $n = 12$ and degree $m = 12$. The lower spherical harmonics of the standard geopotential model such as GEM-10 are considered to have originated from sources deeper than the Earth's asthenosphere⁶. Therefore, the residual geoid map is appropriate for the discussion of structural anomalies in the lithosphere and asthenosphere. Note that the shortest

Table 1 Average residual geoid in the marginal sea basins of the Philippine Sea

Basin	Water depth (m)	Grid for average computation		Residual* geoid		
		Grid size (deg)	No. of grids	Average and s.d. for the basin $\bar{\Delta h} \pm \sigma(\Delta h)$ (m)	Reference values from undisturbed Western Pacific $\bar{\Delta h}_0$ (m)	Reduced residual geoid $(\bar{\Delta h} - \bar{\Delta h}_0) \pm \sigma(\Delta h)$ (m)
Bonin Trough	>3,000	0.5	13	4.1 ± 1.5	0.6	3.5 ± 1.5
Shikoku Basin	>4,000	1.0	33	1.8 ± 2.1		1.2 ± 2.1
Mariana Trough	>3,000	1.0	16	7.0 ± 2.4		5.7 ± 2.4
Parece Vela Basin	>4,000	1.0	85	1.1 ± 2.7	1.3	-0.2 ± 2.7
Northern West Philippine Basin	>5,000	1.0	77	-3.5 ± 1.4		-4.8 ± 1.4
Okinawa Trough	>1,000	0.5	33	1.5 ± 0.8		5.0 ± 0.8

* Lower harmonics of GEM-10 geoid, expanded up to order $n = 12$ and degree $m = 12$, removed from the total geoid.

wavelength contained in the removed component is 30° , about the same as the width of the Philippine Sea. Physiographic provinces necessary for the following discussion are summarized in Fig. 1c.

Geoidal distribution

As expected, the geoid distribution has many features in common with the gravity field which primarily reflects the bottom topography. In particular, the similarity between the distributions of the residual geoid and the free-air gravity anomaly⁷ is

remarkable. Conspicuous geoidal depressions exist on the Bonin, Mariana, Yap, Palau, Philippine and Ryukyu trenches. Geoidal bulges, corresponding to outer highs of topography and gravity seaward of the trenches⁸ are well developed along the Bonin, Mariana and Philippine trenches. All major topographic features in the Philippine Sea have geoidal signatures. The Nankai Trough is associated with a notable geoidal low. The north-south trending South Honshu (West Mariana and Iwo-jima) and Kyushu-Palau ridges are seen as disturbances to the generally smoothly varying geoid of the Philippine Sea;

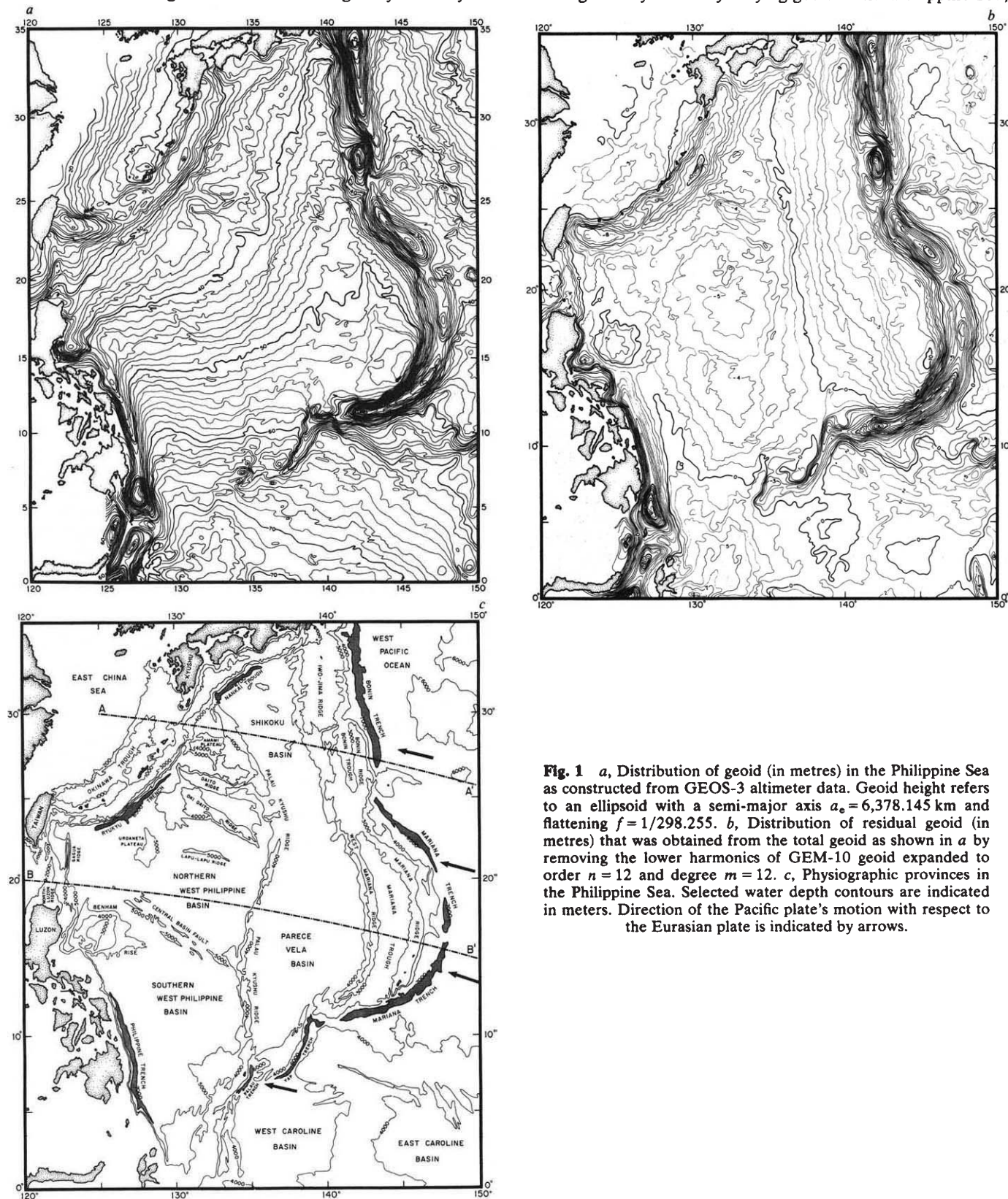


Fig. 1 *a*, Distribution of geoid (in metres) in the Philippine Sea as constructed from GEOS-3 altimeter data. Geoid height refers to an ellipsoid with a semi-major axis $a_e = 6,378.145$ km and flattening $f = 1/298.255$. *b*, Distribution of residual geoid (in metres) that was obtained from the total geoid as shown in *a* by removing the lower harmonics of GEM-10 geoid expanded to order $n = 12$ and degree $m = 12$. *c*, Physiographic provinces in the Philippine Sea. Selected water depth contours are indicated in meters. Direction of the Pacific plate's motion with respect to the Eurasian plate is indicated by arrows.

so are the Daito, Oki-Daito, Lapu-Lapu, Gagua and the Central Basin ridges that are oblique (except Gagua that is parallel) to the trend of the South-Honshu and Kyushu-Palau ridges. Over the Benham Rise and the Urdaneta Plateau, the geoidal anomalies reflect their topography, which is generally smooth, flat and shallower than the surrounding oceanic floors⁹. Quasi-circular negative geoidal anomalies are observed off Taiwan and Luzon where the Ryukyu and the Philippine trenches abut, respectively, against the Gagua Ridge and its extension.

Geoidal height-age relationship

The distribution of the residual geoid in the Philippine Sea (Fig. 1b) shows that the Philippine Sea basin's geoidal height decreases as the basin's age increases. Over the active Mariana Trough (Quaternary), the residual geoid is the highest in spite of its local topographic depression. The geoid is lower in the inactive Shikoku (early Miocene-middle Miocene) and Parece Vela (early Oligocene-middle Miocene) basins. A further decrease of the geoid from the Parece Vela Basin to the Northern West Philippine Basin (Palaeocene-Eocene) seems to reflect the difference in age as well as the mean water depth, which increases substantially from east to west across the Palau-Kyushu Ridge. It can readily be inferred that the source of the residual geoid anomaly, localized in the lithosphere and asthenosphere, is closely related to the mechanism of basin formation.

To substantiate the argument, the Philippine Sea basins' average residual geoid heights were calculated (Table 1). The basins were defined by water depth, and were divided into a set of grids of equal area. Then, at each grid point, an expected geoidal height value was computed by the weighted average method³. A GEOS 3 autocovariance was used as a weight and the range of data collection was set to 1.5 arc deg. This implies that all adjusted GEOS 3 sea-surface height values, within 167 km of a grid point, were used to estimate the expected residual geoid height at that point. The basin's average is simply an arithmetic mean of these expected values.

The average residual geoid heights of the Philippine Sea basins were then normalized against that of the Pacific Ocean floor off the Bonin and Mariana trenches. Formation of marginal sea basin is causatively closely related to the oceanic lithosphere's subduction under the arc-trench system. It is postulated that the marginal sea basin's residual geoid anomaly is meaningful only when compared with the residual geoid over the normal oceanic floor where it is free from the effect of bending due to subduction. Average residual geoid height of the Western Pacific margin was calculated with the aid of $1^\circ \times 1^\circ$ grids "in the region 400 to 500 km seaward of the trench beyond the outer rises in topography where the ocean floor is regionally undisturbed and of normal depth for its age"¹⁰. The results obtained from the latitude range from 20°N to 35°N and from 10°N to 25°N were subtracted, respectively, from the Bonin Trough-Shikoku Basin data and the Mariana Trough-Parece Vela Basin-Northern West Philippine Basin data as shown in Table 1. For the data in the Okinawa Trough, that of the Northern West Philippine Basin was used as a reference.

Figure 2 illustrates the reduced residual geoid heights of the Philippine Sea basins plotted against their ages determined from dating of the oldest sediments from DSDP Legs 6 (ref. 11), 31 (ref. 12), 58 (ref. 13), 59 (ref. 14) and 60 (ref. 15) drillings and the ages estimated from the basin's magnetic lineations^{9,16,17}. The data indicate a linear relationship with a slope of 0.26 m Myr^{-1} . This is a somewhat higher slope than observed on the mid-oceanic ridges, which ranges from 0.11 to 0.19 m Myr^{-1} (W. Haxby, personal communication; see also ref. 18). As McAdoo pointed out¹⁹, the geoidal anomaly over the marginal sea basins includes the effect of a dense subducting lithosphere. In the Bonin and Mariana island arc systems, the dip of Benioff-Wadati zone is among the steepest (70° – 90° in thrust angle) and its downthrust the deepest (500–700 km)²⁰ of the world's subduction systems²¹. The result is a broad and smoothly varying geoidal anomaly with its maximum located

immediately landward of the fore-arc axis. Therefore, it can be assumed that a substantial part of the anomaly due to the subducting lithosphere has already been removed as lower harmonics of GEM-10 geopotential. The anomaly's remaining component, if any, being subtracted from the observation, the slope of the geoid height-age relationship may become indistinguishable from the values exhibited by the mid-oceanic ridges.

Back-arc opening model

The Philippine Sea basin's geoid height-age relationship as summarized in Fig. 2 seems to imply that the marginal sea basin's mode of cooling after its formation is similar to what the oceanic lithosphere experiences in the vicinity of the mid-oceanic ridge. The steeper slope of the linear relationship for the marginal sea basin indicates that the cooling is faster for the lithosphere of marginal sea. This is conceivable because, compared with the oceanic lithosphere created at the mid-oceanic ridge, marginal sea basin's lithosphere has a much smaller extent and in contact with adjacent lithospheres that are generally older, and hence colder, than the newly created lithosphere of the marginal basin.

The data seem to be consistent with the theory of back-arc opening^{22,23}. According to Karig²³, the basins of the Philippine Sea were created by crustal extension. After the formation of the West Philippine Basin, the Shikoku and Parece Vela basins opened and these were, in turn, followed by the opening of the Mariana Trough. An opening occurs by separating a remnant arc, formerly an active arc-trench system, from a continuously active arc-trench system that migrates seaward as spreading of the marginal sea proceeds. The opening is not a continuous

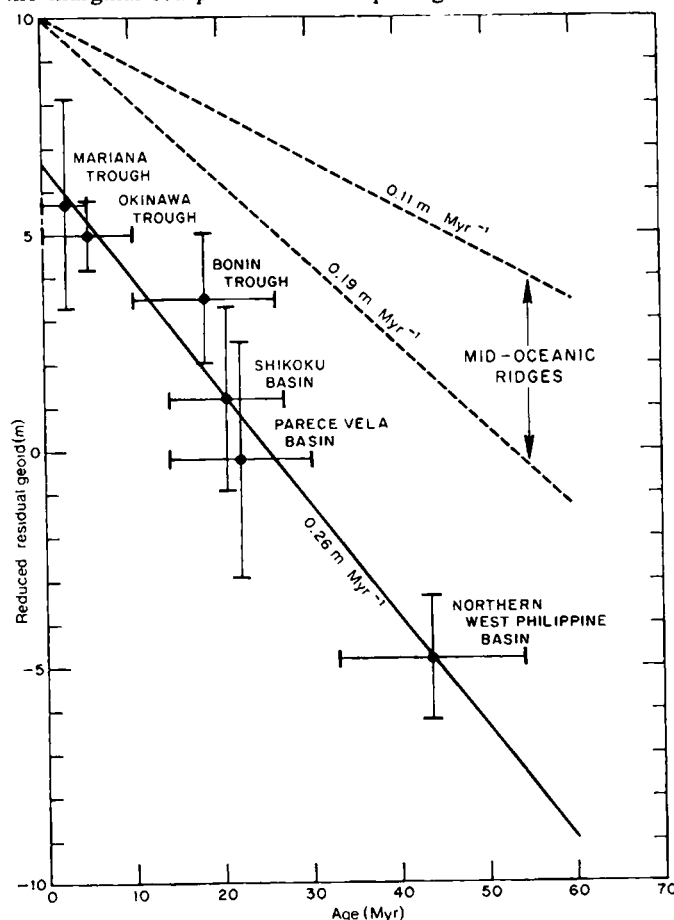


Fig. 2 Average reduced residual geoid height-age relationship in marginal sea basins of the Philippine Sea. Ranges of uncertainties (error bars) of each datum point are: for geoid height, standard deviation as shown in Table 1; for age, ranges of values as indicated by either direct measurement on DSDP samples or inference from magnetic lineations. Ages of the Bonin and the Okinawa troughs are uncertain. Based on the arguments of structural geology of these areas^{27,31}, the ages of Miocene (10–26 Myr) and Pliocene-Pleistocene (0–10 Myr) were assigned, respectively, to the openings of the Bonin and the Okinawa troughs.

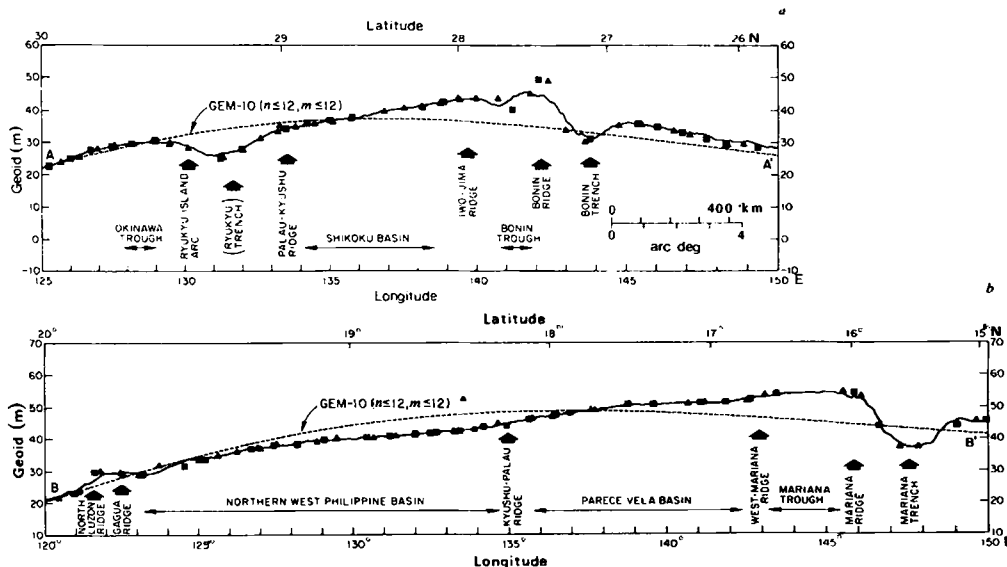


Fig. 3 Profiles of geoid along traverses across the Philippine Sea; *a*, a Bonin Trough-Shikoku Basin-Okinawa Trough traverse; *b*, a Mariana Trough-Parece Vela Basin-Northern West Philippine Basin traverse. Solid line, a profile of expected values calculated as a weighted average of data distributed within 167 km of the traverse. Dashed line, lower harmonics of GEM-10 geoid expanded to order $n = 12$ and degree $m = 12$. Adjusted GEOS-3 track data: ■, ascending; ▲, descending at intersection with the traverse.

process, but two successive openings are separated by a period of quiescence spanning 10–15 Myr. Thus, the Palau-Kyushu Ridge is a remnant arc associated with the openings of the Shikoku and Parece Vela basins and the West Mariana Ridge with that of the Mariana Trough.

Karig²³ modelled the crustal extension of marginal sea as caused by a thermal diapir, supposedly generated at the frictional upper surface of the subducting lithosphere. That the sinking slab's upper surface is the locus of thermal diapir generation has since been questioned²⁴. An equally important problem is whether the crustal extension due to intrusion of thermal diapir is a cause or an effect of the marginal sea's opening. In mid-oceanic ridge areas, the regional tectonic stress field is known to be tensional and the crustal accretion is believed to be due to passive upwelling of basaltic magma that is undoubtedly generated in the upper mantle. In marginal sea areas that are the zones of plate convergence, the tectonic stress is generally compressional. Recent studies tend to indicate, however, that a change of regional tectonic stress pattern from compressional to tensional²¹, or at least a release of compressional stress²⁵, is necessary for a marginal sea to open. It could be safely assumed that the mode of ascent of the thermal diapir in the marginal sea basin is similar to the passive upwelling of magma in the mid-oceanic ridges. In fact, an active spreading centre, comparable with that in the mid-oceanic ridges, has been identified and located in active marginal sea basins such as the Mariana Trough²⁶.

Among the basins of the Philippine Sea that were formed by back-arc opening, the Bonin Trough is anomalous in its mode of formation²⁷. The Bonin Trough is situated seaward of the Iwo-jima Ridge (active fore-arc) that is a complex of *en-echelon* ridge and trough systems with signs indicating nascent rifting. The Bonin Trough's frontal arc is the Bonin Ridge, an uplifted trench slope break. The Bonin Trough itself has no evidence of recent activity, but there is no doubt about its crustal extensional origin. Thus, the Bonin Trough cannot be classified as a back-arc basin, but is better termed a fore-arc basin. In spite of this fact, the Bonin Trough's geoid height-age datum appears to be consistent with the data of back-arc basins (Fig. 2). We conclude that the linear relationship between the reduced residual geoid and the age holds for the marginal sea basin of extensional origin, irrespective of whether the basin is situated fore-arc or back-arc.

Geoidal slope over marginal sea basins

Another remarkable feature of the Philippine Sea's geoidal distribution (Fig. 1*a, b*) is a uniform horizontal space gradient of the geoid over the marginal sea basins. This is best illustrated by the profiles of geoid along the traverses across the sea (Fig. 3*a, b*). Recognizing the causative relationship between the

marginal basin formation and the subduction of oceanic lithosphere, the traverses were chosen in the direction parallel to that of the Pacific plate's convergence to the Philippine Sea²⁸. The profiles show that both the total and the residual geoids' horizontal space gradients are uniform and tilt eastward on the Shikoku and Parece Vela basins. Part of the eastward tilt of the geoid is undoubtedly due to the water depth, which increases to the west in both the Shikoku and Parece Vela basins. The eastward tilt of the geoid could be explained most easily if the openings of the Shikoku and Parece Vela basins were unilateral, with their spreading centres migrated eastward with the advancing fore-arc. Then, it becomes likely that the westward decreasing geoid height is due to the age of the basin floor which increases progressively from east to west.

There is evidence, however, that both basins were formed by bilateral spreading. Based on magnetic lineation and bottom topography, Watts and Weissel¹⁶ traced a N-S trending remnant spreading centre in the middle of the Shikoku Basin. The data are highly disturbed in the eastern half of the basin, and the authors retained a possibility that the basin's mode of opening could be unilateral. A later study¹² confirmed a bilateral spreading as a more likely cause of the Shikoku Basin's opening. In the Parece Vela Basin, Mrozowski and Hayes¹⁷ located the N-S trending Parece Vela Rift and identified it as the approximate location of a remnant two-limbed spreading centre, indicating a similarity of structure and history with the Shikoku Basin.

Although a bilateral spreading in the Shikoku and Parece Vela basins is thus undeniable, the geoidal distribution of these basins shows no sign of geoidal height-age relationship within the basins. The geoid's slope is constant despite the age difference of 8–13 Myr between the centre and the margins of these basins. Analogy with the mid-oceanic spreading centre predicts a decrease of geoid of 1–2 m from the basin's centre to the margins. This is a detectable amount in the light of the nominal precision of 50 cm for the GEOS 3 altimeter measurement. There are some structural anomalies peculiar to these basins. The remnant spreading centre of the Shikoku Basin is not as morphologically prominent as the active spreading centre of the mid-oceanic ridge. In the Parece Vela Basin, the remnant spreading centre is identified as a rift rather than a ridge. In both basins, the sediment thickness increases substantially from the basin's centre towards the margins. A combined effect of these may be large enough to offset the effect of the ageing lithosphere.

In the West Philippine Basin, the NW-SE trending Central Basin Fault was believed to be a remnant two-limbed spreading centre. A recent study⁹ showed that its exact position does not precisely coincide with the Central Basin Fault. Nevertheless, the generally NNE-SSW trending spreading direction was

inferred from magnetic lineation. Figure 1b shows a symmetric distribution of residual geoid with respect to the West Philippine Basin's remnant ridge crest. In its vicinity, a decrease of the geoid for the age increment of 10 Myr is 2 and 1 m, respectively, to the NNE and SSW directions. The slope is steeper for the former, but the data for both limbs agree with that exhibited by the active mid-oceanic ridges.

Corner flow model

The eastward tilt of the total geoid over the entire Philippine Sea and that of the residual geoid over the basins east of the Palau-Kyushu Ridge (Fig. 3a, b) seem to lack a correlation with the basin's lithospheric structure. They must have originated from some causes in the asthenosphere underneath or deeper. As one such cause, a corner flow looks most pertinent. Corner flow is a flow of visco-elastic material in the asthenosphere induced by the subducting oceanic lithosphere. Of the two flow systems induced along the sinking slab's upper and lower sides, that on the upper side (landward) is the more important in the present discussion. McKenzie²⁹ showed that the maximum stress heating of the viscous flow is localized in the vicinity of the island arc, directly beneath the marginal sea's lithosphere. This suggests that the thermal diapir, as a cause of marginal sea opening, is generated in the asthenosphere where the corner flow's viscous stress heating is a maximum. Thus, the corner flow model, an apparent alternative to the thermal diapir model, can actually be compatible with, as well as complementary to, it. Rheology of mantle material is, however, generally not well known. In McKenzie's calculation, the asthenosphere's viscous flow was assumed to be newtonian. It was recently shown³⁰ that the regional geoid depends strongly on the power of non-newtonian constitutive equation ($\dot{\epsilon} = \eta \sigma^n$; $\eta = \text{constant}$) which gives the strain rate ($\dot{\epsilon}$) as a function of

stress (σ). McAdoo's model calculation with a sinking slab's dip of 60° showed that the exponent (n) of the constituent equation must be >3 for the slope of geoid over the marginal sea basin to tilt towards the fore-arc. Apparently, simulations for a sinking slab with much steeper thrust angles are needed to compare the numerical result with the data obtained from the Philippine Sea. The slope of the total geoid ranging from 0.8 to 1.6 m per 100 km over the Philippine Sea basin should impose an effective constraint on the rheological property of the model. It is not clear whether the corner flow is restricted within the mobile asthenosphere or, as the depth of deep-focused earthquakes indicates, penetrating into the Earth's deep interior. If the corner flow's convective cell is shallow, the slope of the residual geoid will be more appropriate for the theory.

Conclusion

In addition to the bottom topography, marine geology, magnetic anomaly and heat flow, the geoid derived from satellite altimeter data has proved useful in studying the marginal sea's history. The idea of comparing the geoid of marginal sea basins with that of the oceanic floor, where the lithosphere is in a pre-subduction state and is undisturbed from the bending effect, is probably the most important. Data of bottom topography and heat flow processed in a similar fashion will serve to develop a systematic theory of marginal sea formation. This study will be extended to other marginal seas in the Western Pacific Ocean.

The original draft of this article has been revised substantially by constructive criticisms given by Dr D. E. Karig of Cornell University. This work was supported by the Office of Naval Research under contract N00014-80-C0098-B. Lamont-Doherty Geological Observatory Contribution No. 3666.

Received 13 January; accepted 16 June 1982.

1. Karig, D. E. *Bull. geol. Soc. Am.* **82**, 323-344 (1971).
2. Rapp, R. H. *J. geophys. Res.* **84**, 3784-3792 (1979).
3. Horai, K. *J. geophys. Res.* (in the press).
4. Watts, A. B. & Daly, S. F. *A. Rev. Earth planet. Sci.* **9**, 415-448 (1981).
5. Lerch, F. J., Klosko, S. M., Laubscher, R. E. & Wagner, C. A. *J. geophys. Res.* **84**, 3897-3916 (1979).
6. Kahn, M. S. *Geophys. J.R. astr. Soc.* **48**, 197-209 (1977).
7. Watts, A. B., Bodine, J. H. & Bowin, C. O. *A Geophysical Atlas of the East and Southeast Asian Seas: Free Air Gravity Field* (Geological Society of America, MC-25, Boulder, 1978).
8. Watts, A. B. & Talwani, M. *Geophys. J.R. astr. Soc.* **36**, 57-90 (1974).
9. Mrozowski, C. L., Lewis, S. D. & Hayes, D. E. *Tectonophysics* **82**, 1-24 (1982).
10. Watts, A. B., Talwani, M. & Cochran, J. R. *Geophys. Monogr.* **19**, 17-34 (1976).
11. Fisher, G. A. *et al. Init. Rep. DSDP Leg 6*, 293-387 (1971).
12. Karig, D. E. *et al. Init. Rep. DSDP Leg 31*, 25-350 (1975).
13. Klein, G. de V. *et al. Init. Rep. DSDP Leg 58*, 21-545 (1980).
14. Kroenke, L. *et al. Init. Rep. DSDP Leg 59*, 21-483 (1981).
15. Hussong, D. M. *et al. Init. Rep. DSDP Leg 60*, 101-249 (1982).

16. Watts, A. B. & Weisell, J. K. *Earth planet. Sci. Lett.* **25**, 239-250 (1975).
17. Mrozowski, C. L. & Hayes, D. E. *Earth planet. Sci. Lett.* **46**, 49-67 (1979).
18. Haxby, W. F. & Turcotte, D. L. *J. geophys. Res.* **83**, 5473-5478 (1978).
19. McAdoo, D. C. *J. geophys. Res.* **86**, 6073-6090 (1981).
20. Katsumata, M. & Sykes, L. R. *J. geophys. Res.* **74**, 5923-5984 (1969).
21. Uyeda, A. & Kanamori, H. *J. geophys. Res.* **84**, 1049-1061 (1979).
22. Packham, G. H. & Falvey, D. A. *Tectonophysics* **11**, 79-109 (1971).
23. Karig, D. E. *J. geophys. Res.* **76**, 2542-2561 (1971).
24. Sleep, N. & Toksoz, M. N. *Nature* **233**, 548-550 (1971).
25. Scholz, C. H., Barazangi, M. & Sbar, M. L. *Bull. geol. Soc. Am.* **82**, 2979-2990 (1971).
26. Karig, D. E., Anderson, R. N. & Bibee, L. D. *J. geophys. Res.* **83**, 1213-1226 (1978).
27. Karig, D. E. & Moore, G. F. *Tectonophysics* **27**, 97-118 (1975).
28. Minster, J. B., Jordan, T. H., Molnar, P. & Heines, E. *Geophys. J.R. astr. Soc.* **36**, 541-576 (1974).
29. McKenzie, D. P. *Geophys. J.R. astr. Soc.* **18**, 1-32 (1969).
30. McAdoo, D. C. *On the Compensation of Geoid Anomalies due to Subducting Slabs* (Technical memorandum TM82157, Goddard Space Flight Center, Greenbelt, 1981).
31. Herman, B. M., Anderson, R. N. & Truchan, M. in *Am. Ass. petrol. Geol. Mem.* **29**, 199-208 (1979).

A demonstration of ocean acoustic tomography

The Ocean Tomography Group: D. Behringer*, T. Birdsall†, M. Brown‡, B. Cornuelle§, R. Heinmiller§, R. Knox‡, K. Metzger†, W. Munk‡, J. Spiesberger‡, R. Spindel||, D. Webb||, P. Worcester‡ & C. Wunsch§

* Atlantic Oceanic and Meteorological Laboratory of NOAA, Miami, Florida 33149, USA

† Cooley Electronics Laboratory, University of Michigan, Ann Arbor, Michigan 48109, USA

‡ Scripps Institution of Oceanography, IGPP, La Jolla, California 92093, USA

§ Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

|| Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

Over the past decade oceanographers have become increasingly aware of an intense and compact ocean 'mesoscale' eddy structure (the ocean weather) that is superimposed on a generally sluggish large-scale circulation (the ocean climate). Traditional ship-based observing systems are not adequate for monitoring the ocean at mesoscale resolution. A 1981 experiment mapped the waters within a 300 × 300 km square south-west of Bermuda, using a peripheral array of moored midwater acoustic sources and receivers. The variable acoustic travel times between all source-receiver pairs were used to construct the three-dimensional (time-variable) eddy fields, using inverse theory. Preliminary results from inversions are consistent with the shipborne and airborne surveys.

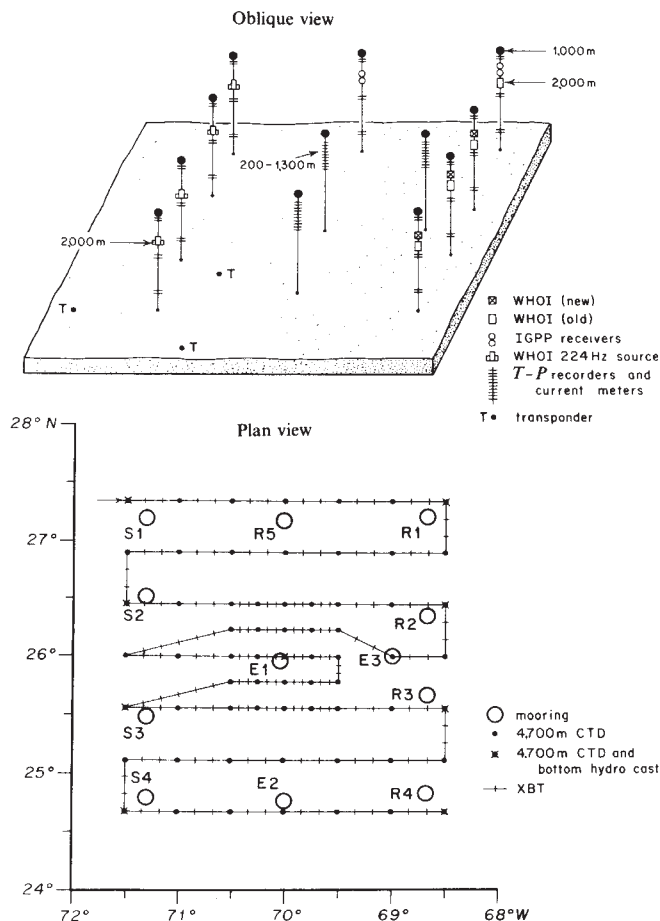


Fig. 1 Geometry of the 1981 Ocean Tomography Experiment. S1 to S4 designate the acoustic source moorings, R1 to R5 the acoustic receiver moorings, and E1 and E2 the environmental moorings with current meters and temperature-pressure (T-P) recorders. E3 is a current meter mooring set by F. Schott. Acoustic moorings also carried T-P recorders. The receivers R1 and R5 consisted of short vertical arrays. Floats at 1,000 m depth with 1,000 lb buoyancy provided taut moorings, thus eliminating wave forces and reducing horizontal excursions. The position of each acoustic source and receiver was monitored by a 10-kHz source (placed just beneath the low-frequency tomographic sources and receivers) which interrogated three bottom transponders T dropped a few kilometres from the mooring (shown for S4). The plan view shows the grid for CTD/XBT surveys in March, May and July. In addition, AXBT drops at the CTD positions were carried out in April and June. The experimental area is over the Hatteras abyssal plain with depths varying between 5,300 and 5,600 m.

THE chief obstacle to understanding the ocean circulation is the overwhelming difficulty in observing it. The ocean is an enormous turbulent fluid fluctuating on all space and time scales. Largely as a result of the MODE expedition¹ in 1973, it has become clear that a 'mesoscale' variability, closely analogous to weather systems in the atmosphere, is superimposed on the large-scale ('climatic') flow field. But the mesoscales are disparate: a few months and several hundred kilometres in the oceans as compared with a few days and several thousand kilometres in the atmosphere.

There has been no perceived national or international requirement to lead towards a large-scale ocean observation network. The cost in ship time is prohibitive (10 full-time vessels for 1 Mm² (1,000 × 1,000 km, about one-third the Mediterranean). Interior *in situ* sensors cannot transmit data by conventional means because the ocean is opaque to electromagnetic radiation. For the same reason remote sensors in space cannot penetrate the ocean interior.

But the ocean is transparent to sound. Acoustic tracking of submarines goes back to World War I and tracking of neutrally buoyant floats from ships or shore stations is now a widely used oceanographic tool. However, there has been no previous

attempt to exploit the properties of the sound field itself as a way of attacking the fundamental observation problem.

The present experiment derives from a proposed scheme² for measuring the field of sound-speed fluctuations within a volume by acoustic transmission through the volume along many diverse paths. It was called 'ocean acoustic tomography', in analogy with the medical X-ray procedure CAT (for computer-assisted tomography). The procedure is best illustrated by reference to the 1981 experiment (Fig. 1). We chose a 300-km square in the same general area occupied by the MODE experiment¹ with a known mesoscale variability (MODE involved six ships and 25 conventional moorings). The 300-km dimension permitted adequate mapping from a ship in three weeks' time for comparison with the acoustic method, and the use of existing acoustic hardware. Sound pulses emitted at each of the four sources S were recorded at each of the five receivers R. Sound-speed in the ocean is predominantly a function of temperature; a cold eddy within the observation square will delay the arrival of any transmission through the eddy.

There is a minimum in sound-speed (Fig. 2) around 1 km depth forming a 'sound channel' which traps some of the acoustic energy. This waveguide, which is the dominant feature of mid-latitude acoustic propagation, is useful here in two ways: (1) acoustic rays that are refracted back towards the axis before intersecting the surface and bottom lose little energy through the boundaries and thus can be detected over several thousand kilometres (megametres), and (2) steep rays which sample the entire water column and generally arrive early can be distinguished from flat late rays which remain nearer the axis, and in this way something can be learned about the depth dependence of the ocean disturbances. A very shallow disturbance affects only the very steep rays with shallow upper turning points; a deep disturbance alters the travel times of both steep and shallow rays.

For each of the 20 source-receiver pairs (Fig. 2 represents one of these) we record roughly 10 multipaths, giving a total of about 200 travel times for each transmission. The number of data points grows geometrically with the product $R \times S$, as compared with the sum $R + S$ for conventional spot measurements and the path integration reduces the noise from local finestructure.

One needs a procedure for converting this information into maps or spectra (or other interesting representations) of the ocean temperature field. Generally the system is underdetermined, and there is an infinity of solutions. 'Inverse theory' can be used to constrain the solution by solving for the smoothest field, or finding the unique solution consistent with *a priori* knowledge of the field covariance. (The procedures of medical tomography are a specialized subset of the available mathematical tools.)

The procedure depends critically on the resolution, identification and stability of multipaths over large ranges. The scheme² proposed in 1979 depended on theoretical estimates, but there was no direct experimental evidence that any one of the three requirements could be met. Positive results, rather

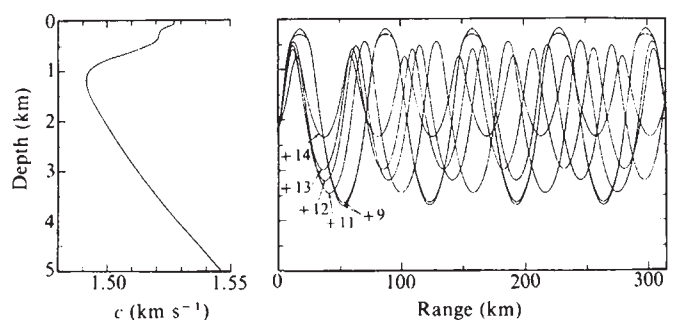


Fig. 2 Ray (multipath) diagram from source S1 to receiver R3 for upward (+) launch angles only. The identifier gives total (upper plus lower) number of ray turning points. The associated measured and predicted arrival structures are plotted in Fig. 3. The range-averaged sound-speed profile from S1 to R3 is shown to the left.

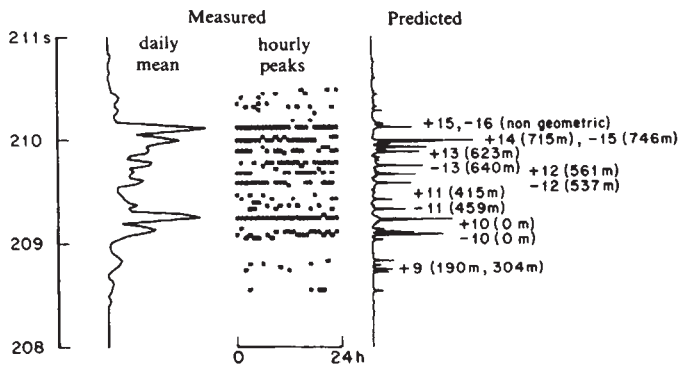


Fig. 3 Comparison of measured arrival pattern with the WKBJ predicted impulse response on year day 64 for S1 to R3 transmissions. The daily mean pattern is formed from 24 hourly transmissions. Travel times for the peak arrivals are shown for each hour. Only peaks with signal-to-noise ratios exceeding 15 dB are plotted. Selected arrivals are identified; for example, +14 (715 m) refers to an upward launch angle with 14 ray turning points, the upper turning point being at 715 m depth. Some of the associated rays are identified in Fig. 2. The non-geometric arrival refers to a ray with a turning point just above the receiver. Surface reflected a non-geometric arrivals are not included in Fig. 2.

better than expected, were subsequently obtained during 48 days of transmissions at 10-min intervals over a 900 km path eastwards from a source moored near Bermuda³. A northward 300-km transmission that was intersected by a southern meander of the Gulf Stream gave similar results⁴. These experiments also provided experience for solving two crucial engineering problems: mooring position keeping and time keeping.

Measurements

The acoustic sources are patterned after the neutrally buoyant SOFAR floats that have been used extensively to track subsurface circulation in the Atlantic^{5,6}. Each source consists of four parallel 1-ft diameter aluminium tubes that are tuned for resonance at 224 Hz and have a bandwidth of 20 Hz. They are driven at the closed end by a flat circular plate of piezoelectric material. A microprocessor controls all source timing (on-off scheduling), including the phase modulation of its 224-Hz carrier. Briefly, the modulation is a 127-digit maximal length shift register sequence⁷. The autocorrelation function of the sequence is sharply peaked and has no sidelobes. Each digit comprises 14 cycles of the carrier and therefore has a duration of 62.5 ms. This is the width of the correlation peak and establishes the multipath resolution of the system. A sequence of 127 digits lasts nearly 8 s, which more than covers the time spread between the first and last arrivals (Fig. 3). A transmission consisting of 24 consecutive sequences lasts nearly 192 s, which is roughly the decorrelation time of the ocean due to internal waves for this experiment. Coherent averaging of the 24-sequence receptions affords signal-to-noise gain. One can think of each transmission as being equivalent to the transmission of 62.5 ms pulses at 8 s intervals repeated 24 times. Coherent summation produces a single intense pulse^{8,9}. The transmitted power level is quite low, about 14 W (186 dB μPa^{-1} at 1 m), for a total of 2,700 J per transmission (about the same energy as in the prolonged grunt of a blue whale heard off the coast of Chile¹⁰) but when this energy is 'compressed' into 62.5 ms pulses it can be recorded at typically 25 dB signal-to-noise ratios, leading to a theoretical travel time precision of 2 ms for a resolved arrival.

The acoustic receivers are microprocessor controlled devices. The arrival time structure of the multipath field between source and receiver is obtained by cross-correlating the coherently averaged incoming signal with a stored replica of the transmitted signal. Some receivers performed the cross-correlation *in situ* and stored only the time of arrival of multipath peaks, while others simply stored signal samples for later on-shore processing. Reception times were synchronized to source transmission times. Transmissions were at hourly intervals (for tidal correc-

tion and to suppress internal wave noise) on every third day only (to conserve batteries).

To provide a time base accurate to better than 10 ms, each source and receiver was equipped with a rubidium atomic frequency standard to make daily measurements of the frequency offset of a considerably less stable (but of lower power) continuously running quartz crystal clock. The frequency offsets are integrated to yield time corrections. It is necessary also to correct the travel times for mooring excursions. These were measured with bottom transponders to a precision of ~ 1.5 m (1 ms for a line-of-sight displacement).

Travel times

The procedure depends critically on the resolution, identification and stability of multipaths over large ranges³. The predicted arrival pattern (Fig. 3) is derived from a generalization to ray theory¹¹, and it matches the measured pattern rather well. This identification, based on travel time, was checked for ray inclination at receivers R1 and R5 which were equipped with vertical arrays¹², and in no instance did the difference between measured and predicted inclination exceed the experimental error.

Figure 4 shows this arrival pattern on a reduced scale over a 100-day period. Multipath peaks are plotted as single points, and only those paths whose signal-to-noise ratio exceeded 15 dB are shown. The relative pattern is stable; some weak late arrivals are surface and bottom reflected and have not been used in the analysis. The 2 s drift of the raw pattern towards

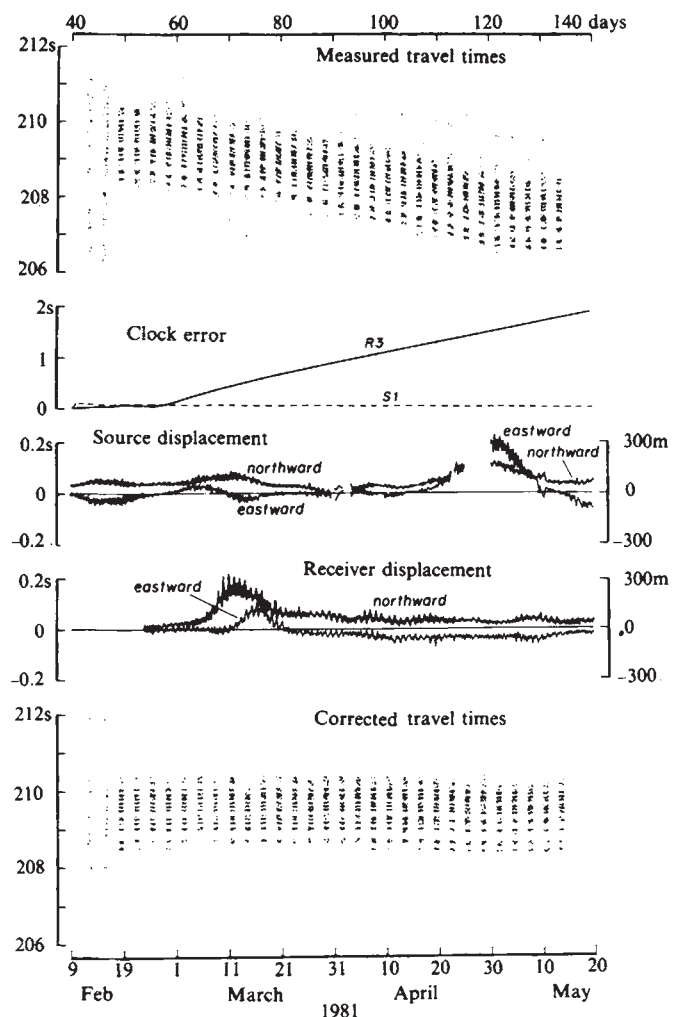


Fig. 4 Multipath arrival structure S1 to R3, before and after corrections for clock drifts and mooring displacements. Transmissions were over 24 h on each third day, starting 18 February (year day 49). The unstructured patterns on days 43 and 46 (before transmission) show processed noise. Gaps in the source displacement are due to missing data.

earlier arrival is almost entirely associated with the gain of the receiver clock (as measured by the rubidium standard). Horizontal excursions of the moorings provide a lesser correction. The high frequency excursions can be related to tidal motion. The arrival pattern corrected for clock drift and mooring motion provides the database for the tomographic inversion.

Comparison of tomographic with traditional measurements

There is a basic problem of comparing two fields that are not the same. The CTD (conductivity/temperature/depth) surveys took almost three weeks (during which time the fields changed); the acoustic surveys provided almost synchronous snapshots (one can think of them as surveys conducted at 3,000 knots).

Panels *b*, *e*, *f*, *i* of Fig. 5 show the sound-speed maps from the first two CTD/XBT cruises¹³. Nominal station spacing was 50 km but near the centre the spacing was halved (Fig. 1). During the second cruise 10 additional 1,500-m CTD casts extended the grid 100 km westward. Panel *a* was inferred (using representative *T-S* relations) from AXBT drops on 13–14 April at positions of the CTD stations. (The AXBT's were from P-3 aircraft supported by NAVOCEANO. Standard AN/SSQ36 probes (to 300 m depth) were used, with temperature calibrations (to $\pm 0.15^\circ\text{C}$) as described in ref. 14.) Sound-speed was computed¹⁵ and then contoured using an objective interpolation scheme like that used in MODE¹⁶. The necessary weights were computed using a gaussian correlation function of 100-km scale.

The March map (days 66–85) shows a cold eddy nearly in the centre, evidently embedded in a background gradient with warmer water to the north. The May map (days 120–139) shows the same cold eddy displaced about 200 km to the west weakened, and perhaps split in two. A new cold eddy, smaller than the original, has appeared in the south-eastern corner.

The four central charts are from the tomographic inversion. The sound-speed field is modelled by specifying a covariance function for the perturbations around a basic state. The horizontal covariance function is homogeneous, isotropic and gaussian with a decay scale of 100 km. The vertical structure is made up of empirical orthogonal functions calculated from the MODE dataset¹.

The tomographic maps show a slow westward movement and slight weakening of the central cold eddy, and the appearance of a new cold centre in the south-east corner, consistent with the CTD surveys. A premature failure of receiver R4, followed by R2 and R3, has so far prevented the plotting of the tomographic chart for the time of CTD II, as had been intended. (The failures were the result of a faulty battery batch; the remaining receivers and all sources functioned normally and will eventually permit limited inversions through the full experimental period.) The chart sequence as shown implies a remarkably stable pattern for the first 40 days, followed by rapid changes in the subsequent 20 days. This interpretation of a 'jumpy' eddy movement is consistent with the observed stability in travel time to R1 until day 110, with rapid changes thereafter in the sense indicated by the charts in Fig. 5. Similarly, 'stick charts' of the measured currents at the environmental moorings E1 and E2 are uneventful until day 110 and then change rapidly¹³. The indicated sequence is as follows¹³:

- Days 62–91: A large cold eddy remains nearly stationary at a position north-east of E1 where it was observed during CTD I. Under its influence the E1 currents flow generally to the south-east and the E2 currents flow to the east.
- Days 91–110: The cold eddy moves slowly WSW to a position west of E1, near where it was seen during the AXBT flight on day 103. (This corresponds to the times of the tomographic inversions.) As a result the flow of the shallow E1 current (137 m) turns counterclockwise to the north, and the

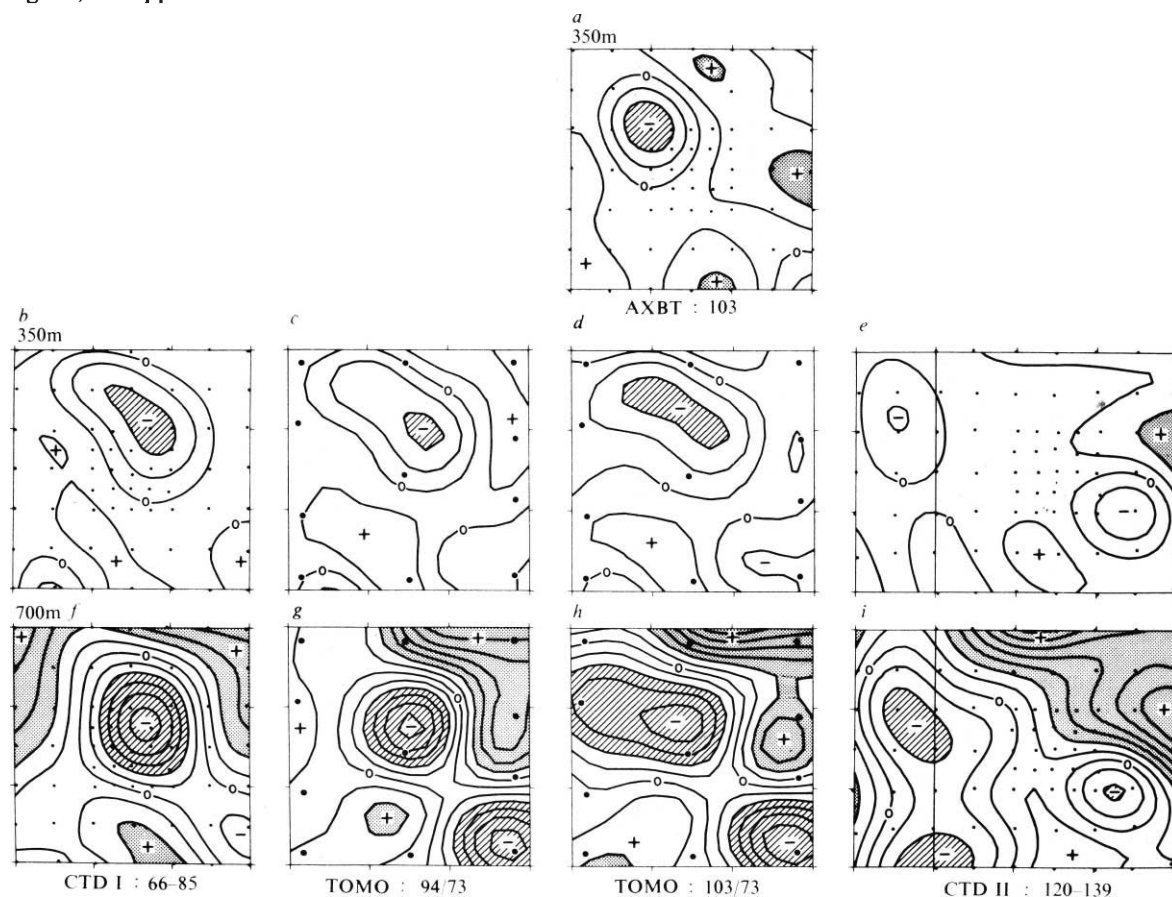


Fig. 5 Mosaic of sound-speed anomaly charts in the 300×300 km square (centred at 26°N , 70°W). *b*, *f* are from the first CTD survey, *e*, *i* from the second CTD survey (CTD II includes an additional 100 km strip to the west); *a* is from the AXBT survey. The four central charts are from the tomographic inversions, assuming travel time error bars of 10 ms. Year days 1981 are indicated below; the tomographic inversions are referenced to the travel time on day 73. *f*–*i* at 700 m depth gives sound-speed departures δC relative to $1,505.9 \text{ m s}^{-1}$; *a*–*e* at 350 m depth relative to $1,521.7 \text{ m s}^{-1}$, with contour interval 1 m s^{-1} ; departures above $+2 \text{ m s}^{-1}$ and below -2 m s^{-1} are shaded. Points give positions of CTD and AXBT casts, and of acoustic and environmental moorings (see Fig. 1 for mooring identification).

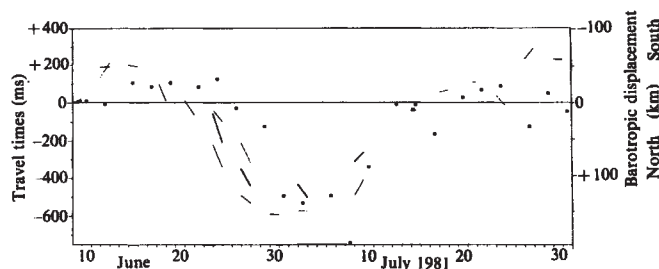


Fig. 6 Relative travel times at hourly intervals for every third day (transmission day) from two sources to north station are portrayed by the sloping line segments; short or missing lines are the result of low signal-to-noise levels. The right-hand scale shows the corresponding barotropic displacement of the Gulf Stream. •, Positions of the North wall of the Gulf Stream inferred from sea-surface temperature measured by NOAA-6 satellite.

flow of the deep (835 m) current meter turns completely around to the west. (The different degrees of turning of the shallow and deep current meters are caused by the north-east tilt of the eddy axis.) The flow of E2 turns slightly clockwise to the ESE.

- Days 110–130: The first cold eddy rapidly moves west to the edge of the experimental box and a second cold eddy appears in the south-east quadrant, where they are seen during CTD II. The flow is generally to the north-west at E1 and south-west at E2.

The above interpretation is supported by the good agreement between the measured currents and those from the dynamic height maps during CTD I and CTD II. We conclude that the tomographic inversions are consistent with the CTD and current meter observations.

The tomographic charts in Fig. 5 are the result of preliminary inversion for a few transmission days, using only a few paths per source–receiver pair. Work is in progress to optimize the procedure and to attempt inversions subsequent to day 103 with the degraded array (unfortunately the number of data points not only grows but also declines geometrically with the number of moorings).

There are two basic problems with the inversion process: initialization and linearization; work on these problems has started. In this preliminary report we follow the simplest initialization procedure: for each source–receiver pair we use only the difference in travel time reckoned from the (known) ocean state on day 73 to compute perturbation fields relative to day 73, and map the combined (initial plus perturbation) fields. For very long time series one would refer the perturbation maps to a climatological mean state which (probably) suffices for ray identification purposes. (Doubtful cases could be tied down with AXBT flights.) Any errors in the assumed mean state will, of course, be retained in the combined maps.

An equivalent (and better) procedure is to use absolute travel time for the known initial (or mean) ocean state to correct the array configuration for a navigational error of ± 0.5 km, and to use absolute travel times in the subsequent inversions. (This involves only 15 independent constants, for example, the x , y positions of eight moorings relative to some arbitrary reference mooring, plus an arbitrary coordinate rotation. The present procedure involves 20 independent constants, for example, the absolute travel times on day 73 for each source–receiver pair.)

Formally, the tomographic procedure is independent of CTD surveys; these are used only to remove navigational errors. It

is expected that the GPS satellite system and sea floor localization procedures¹⁷ will determine the configuration of future arrays to within ± 1 m. At this point a knowledge of the climatological state is required only for ray identification, and the perturbation maps are free from initialization error.

In our linearized inversion procedure we ascribe perturbations in travel time to perturbations in sound-speed along the unperturbed ray paths. But rays emerging from mesoscale eddies have turning depths that may differ sufficiently from the unperturbed turning depths so that a significant part of the travel time perturbation is accumulated outside the eddy¹⁸. Generally the errors are too large to be neglected, yet sufficiently small to suggest an iteration procedure¹⁹. This has not yet been attempted.

Remote stations

The sources were monitored at remote facilities to assure their proper functioning, and to gain experience with long-range transmissions. In this connection we observed a systematic fluctuation in travel time over a northward 2,000-km path across the Gulf Stream (Fig. 6). (On calendar days 52, 55, 58 the receptions from source S2 were poor; on day 61 this source was recovered, found faulty, and another source deployed. No other source anomalies were seen during the remaining six months.) At the midpoint of the 50 days record, the travel time was shortened by 0.8 s. This is the result of a 200-km northward displacement of the Gulf Stream, with the associated decrease in the amount of cold slope water over the path²⁰. (Satellite sea surface temperatures support this interpretation.)

Discussion

The preliminary report of the 1981 experiment indicates that ocean mapping with mesoscale resolution can be performed tomographically over substantial areas. There is, of course, a shortcoming in measuring only the sound-speed field. But in regions of tight T - S relations, or where the disturbances are associated mostly with vertical displacements of the isopycnals, the sound-speed perturbation can be mapped into density perturbations with accuracies adequate for computing geostrophic currents²¹.

Two developments are planned for the future. In the autumn of 1982 we plan to make reciprocal transmissions²² at 300 km ranges and use the difference in the oppositely directed travel times to infer the water velocity along the ray paths (velocity tomography). Subsequently we plan to establish an array of megametre dimensions to measure the variability of an ocean gyre²³. Here the tomographic principle would be used to obtain vertical resolution; in the horizontal directions advantage is taken of the integrating properties of long range sound transmissions, opening the way to true ocean-basin scale measurement programmes.

We thank C. Spofford for discussion and advice. J. Dahlen furnished the T - P recorders. Many people have helped, but we are particularly indebted to Paul Boutin, S. Liberatore, F. Schuler, A. Bradley, J. Kemp and M. Jones of Woods Hole, to F. Dormer, K. Hardy, B. Ma, D. Peckham and R. Truesdale at Scripps, and to B. Grant and S. Johnson at MIT. The work has been supported by the Office of Naval Research and by the NSF. The Atlantic Oceanographic and Meteorological Laboratory of NOAA provided support for three CTD/XBT cruises of the ship *Researcher* and for J.S. The US Navy Oceanographic Office supported the two AXBT flights.

Received 21 June; accepted 7 July 1982.

1. The MODE Group *Deep-Sea Res.* **25**, 859 (1978).
2. Munk, W. & Wunsch, C. *Deep-Sea Res.* **26A**, 123 (1979).
3. Spiesberger, J. L., Spindel, R. C. & Metzger, K. J. *acoust. Soc. Am.* **67**, 2011 (1980).
4. Spindel, R. C. & Spiesberger, J. L. *J. acoust. Soc. Am.* **69**, 982 (1981).
5. Webb, D. C. *Proc. Oceans '77 Conf. Rec. MTS-IEEE* **2**, 44B (1977).
6. Spindel, R. C., Porter, R. P. & Webb, D. C. *IEEE J. Oceanic Engng OE-2*, 331 (1977).
7. Golomb, S. W. *Shift Register Sequences* (Holden-Day, San Francisco, 1967).
8. Spindel, R. C. *Proc. IEEE Electron. Aerospace Convent.* **80** CH, 1587-4AES, 165 (1980).
9. Spindel, R. C. *IEEE Trans. Acoustics, Speech Signal Processing ASSP-27*, 723 (1979).
10. Cummings, W. C. & Thompson, P. O. *J. acoust. Soc. Am.* **50**, 1198 (1971).
11. Brown, M. G. *J. acoust. Soc. Am.* (in the press).
12. Worcester, P. F. *J. acoust. Soc. Am.* **70**, 1743 (1981).

13. Behringer, D. W., Heinmiller, R., Knox, R. & Thomas, G. *NOAA Technical Rep.* (1982).
14. Sessions, M. H. & Barnett, T. P. in *Near-Surface Ocean Experimental Technology Workshop Proc. NORDA*, 125 (1980).
15. Chen, C. T. & Millero, F. J. *J. acoust. Soc. Am.* **62**, 1129 (1977).
16. Bretherton, F. P., Davis, R. E. & Fandry, C. B. *Deep-Sea Res.* **23**, 559 (1976).
17. Spiess, N. F. (Chairman), *Panel on Ocean Bottom Positioning of the Committee of Geodesy* (National Academy of Sciences, Washington, DC, 1982).
18. Mercer, J. A. & Booker, J. R. *J. geophys. Res.* (submitted).
19. Spiesberger, J. L. & Worcester, P. J. *J. acoust. Soc. Am.* (submitted).
20. Spiesberger, J. L. *et al. J. phys. Oceanogr.* (submitted).
21. Worcester, P. F. & Cornuelle, B. *Proc. 2nd Working Conf. on Current Measurements* (in the press).
22. Worcester, P. J. *J. acoust. Soc. Am.* **62**, 895 (1977).
23. Munk, W. & Wunsch, C. *Phil. Trans. R. Soc.* (in the press).

Brain spectrin, a membrane-associated protein related in structure and function to erythrocyte spectrin

Vann Bennett, Jonathan Davis & Walter E. Fowler

Department of Cell Biology and Anatomy, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

An immunoreactive analogue of erythrocyte spectrin has been purified from brain membranes. This protein co-sediments with and cross-links actin filaments, associates with spectrin-binding sites on erythrocyte membranes, and has been visualized by rotary shadowing as an extended, flexible rod. The brain spectrin comprises 3% of the total membrane protein, and may have a major role in mediating linkage of actin to membranes.

MEMBRANE-associated forms of actin have been observed in many types of cells and are thought to participate in cell motility, adherence of cells to surfaces, determination of cell shape and in the organization of cell-surface integral proteins (for review see ref. 1). A direct linkage has been demonstrated between actin and integral membrane proteins in lymphocytes^{2,3}, but the polypeptides required for these associations are unknown. Actin is a component of erythrocyte membranes^{4,5} and its mechanism of association with this simple membrane has been established in detail (reviewed in refs 6, 7). Erythrocyte actin is bound to spectrin in a complex perhaps involving band 4.1 (refs 8–10). Spectrin is an extended rod-like molecule present as a tetramer formed by head-to-head association of dimers¹¹; spectrin dimers consist of two different subunits, arranged as parallel chains. The β -subunit of spectrin is bound to ankyrin^{12–16}, which in turn is associated directly with the cytoplasmic domain of band 3, the major membrane-spanning protein in erythrocytes^{17–19}.

It has been assumed that the association of actin with erythrocyte membranes is not relevant to the arrangement of actin in other cell types, where spectrin could not be detected by radioimmunoassay²⁰. However, polypeptides immunologically related to erythrocyte spectrin^{21,22} and ankyrin^{23,24} have been detected in brain and other tissues. These cross-reacting proteins include soluble and membrane-associated forms. A class of soluble analogues of ankyrin and spectrin have been identified as high-molecular weight microtubule-associated proteins (MAPs)^{22,24}. MAP1, a group of polypeptides of molecular weight (M_r) ~370,000 (370K), contains a component which is related to ankyrin, while MAP2, a polypeptide of M_r 300,000, shares antigenic sites with the α -subunit of spectrin. MAP2 shares with spectrin the ability to bind to actin²⁵ and the shape of an extended rod²⁶, but lacks membrane-binding activity. Here we describe another immunoreactive analogue of spectrin from brain which binds to actin, associates with spectrin-binding sites in erythrocyte membranes, and is an extended rod-like molecule. This spectrin analogue comprises 3% of the total membrane protein, and is a suitable candidate for having a major role in mediating linkage of actin to membranes.

Purification of spectrin analogue from brain membranes

Brain membranes contain a prominent pair of polypeptides of M_r ~260K and 265K which cross-react with affinity-purified antibody against erythrocyte spectrin (Fig. 1). Membrane fractions from kidney, liver and testes also contain polypeptides of similar molecular weight which cross-react with anti-spectrin antibody, but brain membranes were the richest source of these polypeptides (not shown). In liver, cross-reacting polypeptides of M_r 260K and 265K were localized exclusively in the plasma

membrane fraction (Fig. 2). The mitochondria, nuclei and soluble fractions all contain polypeptides unique for each fraction which cross-react with anti-spectrin antibody, but none are of molecular weight 260K or 265K. The significance of multiple cross-reacting polypeptides is unknown, and further studies with each polypeptide are required. The reactions are specific in that there was no cross-reaction of these polypeptides with non-immune antibody (Figs 1, 2) or with antibody against ankyrin or band 3 (not shown). A possible explanation is that functional domains of spectrin are conserved as discrete amino acid sequences in other polypeptides and have been combined with other activities to form proteins distinct from spectrin. Here we present a detailed study of the 260K and 265K cross-reacting polypeptides.

Fractionation of brain membranes by differential centrifugation or sedimentation on gradients of sucrose or sucrose-Ficoll did not yield membranes significantly enriched in these polypeptides, and only myelin lacked the polypeptides (not shown). Therefore, total membranes depleted of myelin were used as starting material for purification procedures. These membranes contained ~75% of the 260K/265K polypeptide present in total brain homogenates. The 260K/265K polypeptides can be solubilized from brain membranes by extraction with 1 M KCl or at low ionic strength with 0.2 mM Na-EGTA (not shown). Extraction with low salt was more selective, and was therefore used to solubilize the protein for purification. The immunoreactive 260K/265K polypeptides were purified to homogeneity by precipitation with ammonium sulphate, gel filtration and finally rate zonal sedimentation on sucrose gradients (Fig. 1).

To examine the cross-reaction between spectrin and the brain protein, ¹²⁵I-labelled brain protein was immunoprecipitated by antibody against spectrin, and this binding was displaced completely by unlabelled spectrin (Fig. 1B). The brain protein displaced binding of ¹²⁵I-labelled spectrin to antibody by a maximum of 20% (Fig. 1C). The limited displacement achieved by the brain protein occurred in the same range of concentration as displacement by erythrocyte spectrin. Thus, only 20% of the antigenic sites recognized by our antibody in erythrocyte spectrin are shared by the brain protein, but those sites which are common between these proteins are well conserved. The fact that the brain and erythrocyte proteins are only partially cross-reactive could explain previous negative results of radioimmunoassay for spectrin²⁰.

The possibility that our antibody was reacting with a minor contaminant which co-migrated with the brain polypeptides was excluded by comparing peptide maps of each brain polypeptide with maps of polypeptides immunoprecipitated by the antibody (Fig. 3). The total brain polypeptides and the fraction immunoprecipitated with anti-spectrin antibody produced almost identical peptide maps. Thus, the antibody recognizes a major constituent of each polypeptide.

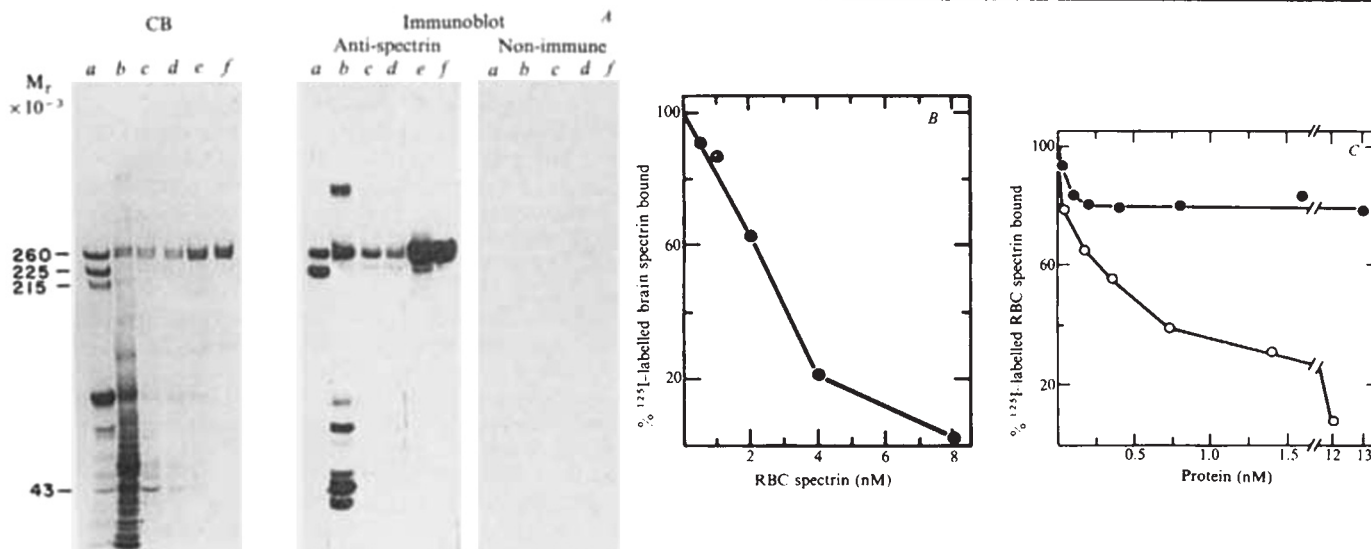


Fig. 1 Purification from brain membranes of a protein which shares antigenic determinants with erythrocyte spectrin. Fractions at various stages of purification were analysed by SDS-polyacrylamide gel electrophoresis on 3.5–17% exponential gradient gels^{22,24} (A). Polypeptides were visualized by staining with Coomassie blue (CB); those cross-reacting with erythrocyte spectrin were identified by electrophoretic transfer from polyacrylamide gels to nitrocellulose paper, followed by overlay of the nitrocellulose with $1 \mu\text{g ml}^{-1}$ affinity-purified rabbit antibody against human erythrocyte spectrin or with non-immune antibody followed by ^{125}I -labelled Protein A^{22,24}. Immunoreactive bands were then detected by autoradiography. Samples are: a, pig erythrocyte ghosts; b, pig brain membranes; c, low ionic strength extract of brain membranes; d, soluble protein recovered from ammonium sulphate precipitation of sample c; e, pooled fractions from gel filtration; f, pure protein from sucrose gradients. Immunoprecipitation of ^{125}I -labelled pure protein (lane f) and ^{125}I -labelled pig erythrocyte spectrin by anti-spectrin antibody were measured as described elsewhere^{22,23}. Binding of ^{125}I -labelled brain protein by anti-spectrin antibody (100 nM IgG) was displaced by unlabelled erythrocyte spectrin (B). C, binding of ^{125}I -labelled erythrocyte spectrin by anti-spectrin antibody (7 nM IgG) was displaced by unlabelled erythrocyte spectrin (○) and partially displaced by unlabelled pure brain protein (●). Frozen grey matter from pig brain (125 g) was homogenized in 800 ml of 0.32 M sucrose, 2 mM Na-EGTA, 200 $\mu\text{g ml}^{-1}$ phenylmethylsulphonyl fluoride (PMSF) (v/v) diisopropylfluorophosphate, 3 $\mu\text{g ml}^{-1}$ pepstatin A, pH 7.0. The 900g supernatant of this suspension was pelleted at 30,000g for 25 min to collect the crude membrane fraction. This fraction was depleted of myelin by centrifugation for 40 min at 30,000g through 0.8 M sucrose. The myelin-depleted membranes were washed successively by centrifugation for 25 min at 30,000g with homogenization buffer, followed by 10 mM sodium phosphate, 2 mM Na-EGTA, 50 $\mu\text{g ml}^{-1}$ PMSF pH 7, and finally 2 mM Na-EGTA, 50 $\mu\text{g ml}^{-1}$ PMSF pH 7. The washed membranes (1.6 g protein) were extracted at 37 °C for 60 min in 800 ml of 0.2 mM Na-EGTA, 0.5 mM dithiothreitol (DTT), 50 $\mu\text{g ml}^{-1}$ PMSF pH 7. The suspension was pelleted at 200,000g for 30 min and the supernatant (120 mg extracted protein) was adjusted to 45% saturation with solid ammonium sulphate. The precipitated proteins were dialysed overnight against 1 M urea, 20 mM glycine, 2 mM Na-EGTA, 1 mM Na₂S₂O₈, 0.2 mM DTT, and centrifuged at 200,000g for 3 h. The supernatant was applied to a Sephacryl S400 gel filtration column (2.6 × 90 cm) equilibrated with urea buffer. Fractions containing the 260K polypeptides eluted near the void volume. The protein was precipitated with ammonium sulphate, dialysed extensively against 0.2 mM Na-EGTA, 0.5 mM Na₂S₂O₈, 0.2 mM DTT pH 7, then centrifuged at 200,000g for 90 min. The supernatant was fractionated on 10–30% linear sucrose gradients in dialysis buffer using a Beckman V Ti 50 rotor (14 h at 36,000 r.p.m.). The gradient fractions with the purest protein (at about a sedimentation rate of 11S) contained 1.2 mg of protein. An additional 2 mg of less pure protein was recovered from adjacent fractions.

A feature which distinguishes spectrin from other high-molecular weight actin-binding proteins, such as filamin and MAP2, is the presence of two distinct polypeptide chains. In this respect, the brain protein resembles spectrin as the 265K and 260K polypeptides are clearly different by peptide mapping (Fig. 3). The two chains co-purify in a 1:1 molar ratio, and both polypeptides co-sediment with actin (see below). The brain protein has a molecular weight of 1,090,000, calculated from preliminary hydrodynamic values ($R_s = 240 \text{ \AA}$, $s_{0,w}^0 = 10.8$, $\bar{v}$ assumed to be 0.73). This protein is therefore a tetramer with two copies of each subunit.

The peptide maps of α - and β -chains of erythrocyte spectrin are dissimilar from either chain of the brain protein (Fig. 3). It is of interest that peptide maps of human and pig erythrocyte spectrin differ by at least 60% and that antibody against human spectrin binds with a lower affinity to pig erythrocyte spectrin (not shown). Erythrocyte spectrin evidently does not have a highly conserved amino acid sequence. The fact that the peptide fingerprints of the brain protein differ from those of erythrocyte spectrin does not rule out functional homology between these proteins or a possible common evolutionary origin. The difference in peptide maps between the immunoprecipitated brain polypeptides and spectrin does, however, exclude the possibility of contamination by erythrocyte spectrin in the brain protein.

Spectrin analogue has similar morphology to erythrocyte spectrin

Erythrocyte spectrin is distinctive in that the tetramer appears in rotary-shadowed preparations as a highly extended rod almost 200 nm long¹¹. In similar preparations, the brain protein also appears as an extended rod with average length 195 nm and width 4–6 nm (estimated without compensating for the

additional mass of platinum) (Fig. 4). Occasionally, the rods open in the middle to reveal two parallel chains. The chains are most often closely apposed or helically entwined along the length of the molecule. Pig erythrocyte spectrin tetramers are slightly shorter with an average length of 175 nm, but otherwise are quite similar to the brain protein (Fig. 4). One significant difference between the two proteins is that, on average, the brain protein is straighter and there are fewer cases of separation of individual chains than observed for erythrocyte spectrin (Fig. 4A; compare with ref. 11). A more rigid conformation for the brain protein tetramer is indicated also by its lower sedimentation coefficient (10.8S) compared with that of the spectrin tetramer (12.5S). The brain protein tetramer was not converted to dimer by warming at 37 °C (not shown) and thus is more stable than tetramers of erythrocyte spectrin. However, a limited conversion of tetramer to dimer was achieved by prolonged exposure to 1 M NaBr (Fig. 4). The dimers produced by this treatment exhibit a R_s of $\sim 110 \text{ \AA}$ and are comparable in length to erythrocyte spectrin dimers.

An important similarity between brain and erythrocyte proteins is that at concentrations less than 2 mg ml^{-1} , oligomers larger than tetramers were observed only rarely. Thus, it is likely that dimers of the brain protein have a polarity and are associated head-to-head, as has been suggested for erythrocyte spectrin^{11,27,28}. Otherwise, with a non-polar or a head-to-tail association, polymerization could occur to form extended filaments.

The spectrin analogue binds to erythrocyte membranes and to F-actin

Important criteria for a spectrin-like protein, in addition to structural similarities, are functions of binding to membranes and association with actin filaments. The brain protein remains

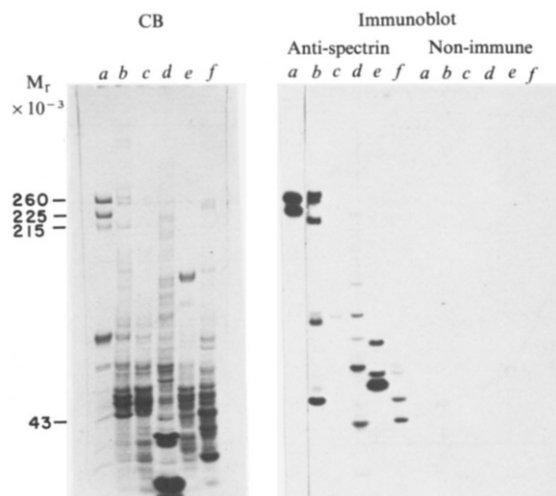


Fig. 2 Polypeptides (265K and 260K) cross-reacting with anti-spectrin antibody are localized in plasma membranes of liver. Plasma membrane sheets from blanch liver were isolated by two steps of upward flotation through 1.42 M sucrose essentially as described elsewhere⁴⁰ except that the buffers contained 7.5 mM sodium phosphate, 1 mM MgCl_2 pH 7.5, and protease inhibitors (200 $\mu\text{g ml}^{-1}$ PMSF and 0.01% (v/v) diisopropylfluorophosphate). Mitochondria were isolated as described elsewhere⁴¹ with the same buffer and protease inhibitors as above. The total microsomal and soluble fractions were isolated by pelleting (200,000g, 45 min) the 25,000g supernatant from liver homogenized in 0.25 M sucrose, 7.5 mM sodium phosphate, 1 mM MgCl_2 pH 7.5 and protease inhibitors. The microsomal pellet was washed once in 7.5 mM sodium phosphate, 1 mM Na-EGTA pH 7.5. Nuclear membranes were prepared essentially by the procedure of Dwyer and Blobel⁴² except that 0.5 mM Na-EGTA and protease inhibitors were included in the buffers, and the nuclei were digested with RNase and DNase. The samples were analysed on SDS-polyacrylamide gels; polypeptides were visualized by staining with Coomassie blue (CB; left) and immunoreactive polypeptides (right) identified as described in Fig. 1 legend. a, Erythrocyte ghosts; b, plasma membranes; c, microsomal membranes; d, nuclear envelopes; e, mitochondria; f, soluble proteins.

associated with membranes during repeated washes, and presumably is involved in a stable linkage with some membrane component(s). The brain protein also contains a binding site for erythrocyte ankyrin, as it displaces binding of ^{125}I -labelled erythrocyte spectrin to inside-out spectrin-depleted vesicles from erythrocyte membranes (Fig. 5). This assay measures association of spectrin with ankyrin, as spectrin binding to the membrane is displaced entirely by a spectrin-binding proteolytic fragment of ankyrin²⁹ and by antibody against this fragment of ankyrin¹². The brain protein is less active than erythrocyte spectrin, with 50% displacement at 60 nM compared with 30 nM for erythrocyte spectrin dimer. ^{125}I -labelled brain protein associates with erythrocyte membrane vesicles (not shown); this binding is blocked by unlabelled erythrocyte spectrin with 50% displacement at 10–15 nM (Fig. 5). The brain protein and erythrocyte spectrin thus bind to the same membrane site but erythrocyte spectrin binds with approximately twofold higher affinity. The lower affinity of the brain protein may be explained by a difference in its ankyrin-binding site, and/or may be a consequence of steric restrictions resulting from its more rigid conformation.

The brain protein associates with actin filaments, as demonstrated by co-sedimentation with actin through sucrose barrier gradients (Fig. 6). Over 60% of the brain protein was pelleted in the presence of actin, compared with <7% in the absence of actin, in conditions of 60 nM brain protein and 5 μM actin monomer. The brain protein greatly increased the low-shear apparent viscosity of F-actin, with an apparent gel point in the range of 60 nM (Fig. 6). Pig erythrocyte spectrin (tetramer/dimer ratio of ~3) exhibited a comparable activity (Fig. 6). The gel point for the brain protein varied with the actin preparation used and was as low as 20 nM when tested with actin purified further by gel filtration (V. Fowler, unpublished data).

Previous studies of cross-linking of actin filaments by human erythrocyte spectrin have noted the requirement for band 4.1 (refs 9, 10). We have also observed that human spectrin alone is not active in cross-linking actin, but that pig erythrocyte spectrin purified in the same way and compared in the same assay is not dependent on band 4.1 (V. Fowler and V.B., unpublished data). The results with pig spectrin agree with those of Brenner and Korn⁸ who observed that pure sheep erythrocyte spectrin was also capable of cross-linking actin

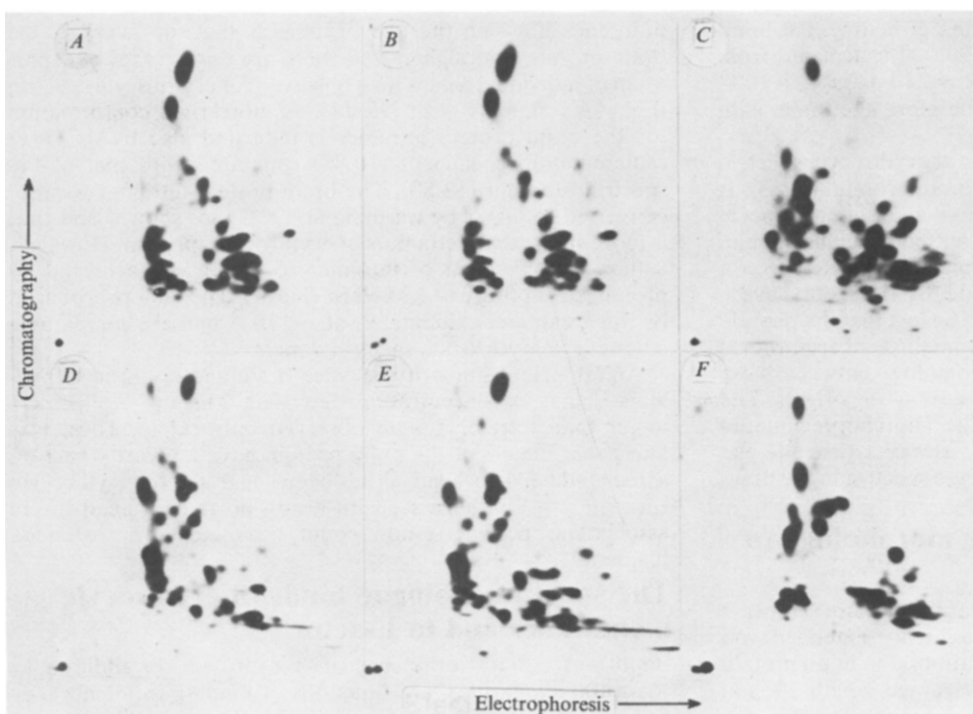
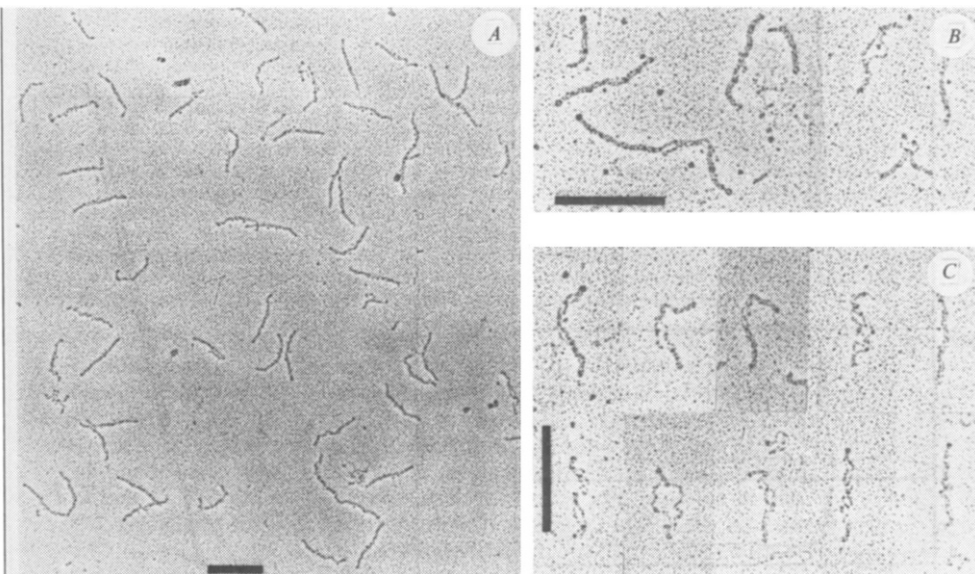


Fig. 3 Two-dimensional maps of ^{125}I -labelled α -chymotryptic peptides from α - (A–C) and β - (D–F) subunits of brain spectrin (A, D), brain spectrin immunoprecipitated by anti-erythrocyte spectrin antibody (B, E) and pig erythrocyte spectrin (C, F). ^{125}I -labelled brain spectrin (Fig. 1; 4×10^7 c.p.m.), ^{125}I -labelled brain spectrin immunoprecipitated by antibody against erythrocyte spectrin (Fig. 1; 2×10^6 c.p.m.) and pig erythrocyte spectrin (3×10^7 c.p.m.) were mixed with 2 μg unlabelled brain spectrin or erythrocyte spectrin and then electrophoresed on a SDS-polyacrylamide gel (see Fig. 1 legend). The gel was stained lightly with Coomassie blue, and equilibrated with 50 mM ammonium acetate, 1 mM sodium azide, pH 7. The protein bands were excised, suspended in 0.5 ml ammonium acetate buffer and incubated at 37 °C for 20 h with two additions of 25 μg α -chymotrypsin. The digests were lyophilized, and analysed by electrophoresis and chromatography as described elsewhere^{22,43}.

Fig. 4 Electron micrographs of rotary-shadowed brain spectrin (A, B and top row of C) and pig erythrocyte spectrin (bottom row of C). A, representative field of unfixed brain spectrin molecules which were diluted to $10 \mu\text{g ml}^{-1}$ in 100 mM ammonium formate (pH 7), mixed with 1/3 vol of glycerol, sprayed on to mica, dried under vacuum at room temperature and rotary-shadowed with platinum at an angle of 1 in 10 (refs 11, 44). Scale bar, 200 nm. B, brain spectrin molecules from a different less pure preparation, which was treated with 1 M NaBr for several days before dialysis against 100 mM ammonium formate and rotary shadowing as described for A; the field on the left includes an 'octamer' and two 'dimers', as well as two 'tetramers'. The two-stranded character of the molecules is seen more clearly in the field on the right and in C. Scale bar, 200 nm. C, selected brain spectrin molecules are shown in the top row and pig erythrocyte spectrin tetramers in the bottom row. These molecules were selected to demonstrate structural similarities between the two spectrin species, such as their two-stranded, double-helical nature. Scale bar, 200 nm.



filaments. Overloaded gels of brain spectrin exhibited contaminating polypeptides, but each polypeptide contributed no more than 1–2% of the total protein. Furthermore, the contaminating polypeptides were not specifically adsorbed by F-actin (Fig. 5). It is therefore unlikely that additional proteins are required in a 1:1 ratio with spectrin for association of brain spectrin with actin. The possibility that a protein may act in sub-stoichiometric ratios cannot be excluded.

Discussion

Here we have described the purification from brain membranes of an immunoreactive analogue of erythrocyte spectrin which closely resembles spectrin in structure and shares functions of binding to erythrocyte membranes and to actin filaments. The brain analogue of spectrin comprises 3% of the total membrane protein, and it would be surprising if this protein had not been noticed previously. An actin-binding protein containing two polypeptides of molecular weight $\sim 240\text{K}$ and 235K has been purified partially from brain and has been termed brain actin-binding protein³⁰. An axonally transported protein with two polypeptide chains of molecular weight $\sim 250\text{K}$ and 240K has been isolated from the particulate fraction of brain, and demonstrated to bind to actin³¹. This protein was named fodrin and was localized by immunofluorescence in the cortical cytoplasm of cultured fibroblasts, neurones and intestinal epithelial cells. It was concluded that fodrin was not related to spectrin as antibody against fodrin did not form a precipitin reaction with erythrocyte spectrin. However, this result does not exclude cross-reactivity that is partial or of reduced affinity.

It is possible that several groups have discovered the same protein and named it differently. We suggest that, in view of the similarity to erythrocyte spectrin, this protein be referred to as brain spectrin. Polypeptides of similar molecular weight to brain spectrin which cross-react with spectrin antibody are present in liver plasma membranes (Fig. 2) and membranes from other tissues (not shown). These proteins are probably also spectrin analogues and can be named according to the cell or tissue of origin.

The extent of structural homology between erythrocyte spectrin and the brain analogue is of the order of 20% as determined by immunoreactivity (Fig. 1). This is much less than the homology observed in other families of proteins such as actins, calmodulins and intermediate filament polypeptides. It should be kept in mind that spectrin polypeptides are unusually large and a homologous region could have a molecular weight of as

much as 60,000. Furthermore, the extended rod-like conformation of spectrin may be compatible with more variation in amino acid sequence than the folding of a globular protein. It is pertinent in this regard that peptide maps of erythrocyte spectrin from different species are substantially different (see above), even though the proteins are almost identical in morphology. In view of these considerations and the demonstration of functional homology, we suggest that the spectrins are a multigene family of proteins with a common evolutionary origin. Erythrocyte spectrin is probably the most deviant member of

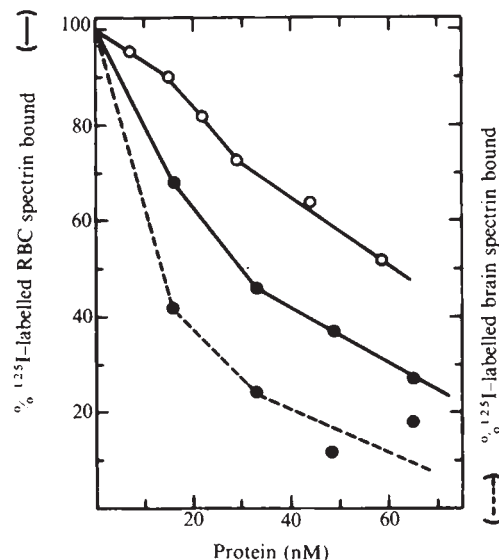


Fig. 5 Displacement of binding of ^{125}I -labelled erythrocyte (—) and brain (---) spectrin to spectrin-depleted erythrocyte membranes by unlabelled erythrocyte (●) and brain (○) spectrin. Binding of ^{125}I -labelled erythrocyte spectrin dimer (1 nM , 2×10^6 c.p.m. pmol^{-1}) and ^{125}I -labelled brain spectrin (1 nM , 8×10^5 c.p.m. pmol^{-1}) to spectrin-depleted inside-out vesicles from human erythrocytes ($20 \mu\text{g}$ membrane protein per assay) was measured as described elsewhere⁴⁵, in the presence of increasing concentrations of unlabelled brain spectrin (○) and pig erythrocyte spectrin dimer (●). The data are expressed as per cent binding in the absence of added proteins, and are corrected for nonspecific binding at each concentration of unlabelled proteins by subtracting values for ^{125}I -labelled proteins that were heat-denatured (10 min at 70°C).

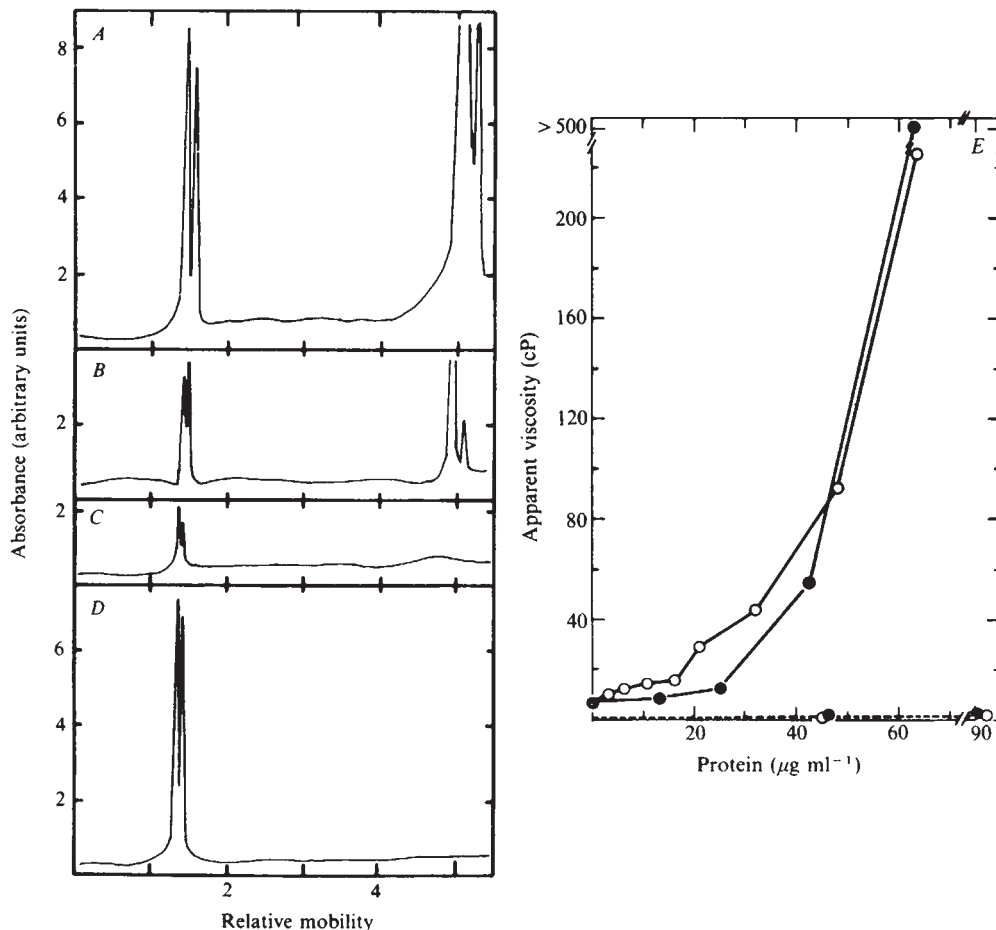


Fig. 6 Brain spectrin co-sediments with F-actin and increases the low-shear viscosity of F-actin. Brain spectrin ($60 \mu\text{g ml}^{-1}$) was incubated (60 min at 30°C) in the presence or absence of rabbit muscle actin⁴⁶ ($200 \mu\text{g ml}^{-1}$) in buffer containing 40 mM PIPES pH 6.8, 20 mM KCl, 1 mM MgCl_2 , 5 mM Na-EGTA. The proteins were layered over 10% (w/v) sucrose in the same buffer and centrifuged at $200,000g$ for 60 min. The pellets were resuspended in the original volume, and equal aliquots of pellets and supernatants analysed by SDS-polyacrylamide gel electrophoresis (see Fig. 1). The polypeptides were visualized with Coomassie blue, and the gels were scanned in a densitometer (left): A, pellet, brain spectrin and actin; B, supernatant, brain spectrin and actin; C, pellet, brain spectrin alone; D, supernatant, brain spectrin alone. E, effect of increasing concentrations of brain spectrin ($\circ$) on the low-shear viscosity of actin (0.8 mg ml^{-1}) was determined in the same buffer with a falling ball viscometer^{10,47}. For comparison, pig erythrocyte spectrin (tetramer/dimer ratio = 3) was tested in the same assay ($\bullet$). In the absence of actin the proteins had very low viscosity (---).

this family, as the functional constraints in these specialized cells are quite different from those in more complex cell types.

The widespread distribution of spectrins implies that functional analogues of ankyrin and band 3 should also be present in other cell types. A polypeptide of $M_r \sim 200,000$ that cross-reacts with antibody against ankyrin is associated with membranes of brain and the plasma membrane fraction of liver (not shown; ref. 32). Immunoreactive forms of band 3, detected with antibody against the cytoplasmic domain of band 3, are also present in the same membrane fractions as ankyrin and spectrin (not shown). It will be essential to purify and characterize these cross-reacting polypeptides to verify possible functions.

Spectrin in association with actin and band 4.1 forms a supportive lattice on the inner surface of the plasma membrane in erythrocytes, and spectrin may also participate in a similar structure in plasma membranes of other cells. Actin is associated with a detergent-insoluble matrix underlying the plasma membrane^{33,34}, and has been identified in association with the inner surface of the plasma membrane in several cases¹. Cells in tissues also have specialized regions of cell-cell contact such as zonula adherens and postsynaptic densities where actin filaments are attached to the membrane. Spectrin may function in these areas to couple actin to the membrane and to form a two-dimensional meshwork adjacent to the membrane analogous to the matrix in erythrocytes.

The spectrins are large enough to be visualized by high resolution electron microscopy of cells and tissues, and would appear as flexible filaments 2–4 nm wide. Filaments having these dimensions and lengths up to 300 nm have been observed in whole mounts of cells by high voltage electron microscopy

and seem to link actin to other cell structures³⁵. Filaments 4 nm wide have also been reported in postsynaptic densities—these were visualized by rapid freezing and rotary shadowing³⁶, and appear to interact with actin filaments as well as intramembrane particles. The narrow filaments do not correspond to previously studied structural proteins and may, at least in part, be composed of spectrin.

The widespread distribution of proteins closely related to erythrocyte spectrin and possibly other erythrocyte proteins suggests it will soon be possible to define in detail at least one generally occurring type of linkage of actin filaments to membranes. This information may explain how integral membrane proteins are organized and directed in their motions and will also be important in understanding how membranes are integrated with the cytoplasm.

Following submission of this article, a protein of similar structural features to brain spectrin has been purified from brain and intestine³⁷. Immunoreactive forms of spectrin and their distribution have also been observed in several tissues³⁸. A brain protein having similar properties to brain spectrin has been isolated³⁹.

This research was supported by grants to V.B. from the NIH (1 RO1 AM29808-1 and a Research Career Development Award) and the Muscular Dystrophy Association. W.E.F. was supported by a postdoctoral fellowship from the Muscular Dystrophy Association. Facilities for electron microscopy were funded in part by grants to Dr Ueli Aepli from the Muscular Dystrophy Association and the NIH (GM2-7765). We thank Drs Ann Hubbard and Larry Gerace for advice on the preparation of plasma membranes and nuclear envelopes from liver. Dr V. Fowler provided muscle actin and advice on measurement of low-shear viscosity.

5. Tilney, L. G. & Detmers, P. J. *Cell Biol.* **66**, 508–520 (1975).

6. Branton, D., Cohen, C. M. & Tyler, J. *Cell* **24**, 24–32 (1981).

7. Bennett, V. J. *cell. Biochem.* **18**, 49–66 (1982).

8. Brenner, S. L. & Korn, E. D. *J. biol. Chem.* **254**, 8620–8627 (1979).

9. Ungewickell, E., Bennett, P. M., Calvert, R., Ohanian, V. & Gratzner, W. *Nature* **280**, 811–816 (1979).

Received 14 April; accepted 11 June 1982.

1. Weighing, R. B. *Meit. Achieve. exp. Path.* **8**, 42–109 (1979).

2. Koch, G. & Smith, M. *Nature* **273**, 274–278 (1978).

3. Flanagan, J. & Koch, G. *Nature* **273**, 278–281 (1978).

4. Ohnishi, T. *J. Biochem.* **52**, 307–308 (1963).

10. Fowler, V. & Taylor, D. L. *J. Cell Biol.* **85**, 361–376 (1980).
11. Shotton, D. M., Burke, B. E. & Branton, D. *J. molec. Biol.* **131**, 303–329 (1979).
12. Bennett, V. & Stenbuck, P. *J. biol. Chem.* **254**, 2533–2541 (1979).
13. Luna, E., Kidd, G. & Branton, D. *J. biol. Chem.* **254**, 2526–2532 (1979).
14. Yu, J. & Goodman, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2340–2344 (1979).
15. Calvert, R., Bennett, P. & Gratzner, W. *Eur. J. Biochem.* **107**, 355–361 (1980).
16. Litman, D., Hsu, C. J. & Marchesi, V. T. *J. Cell Sci.* **42**, 1–22 (1980).
17. Bennett, V. & Stenbuck, P. *Nature* **280**, 468–473 (1979).
18. Bennett, V. & Stenbuck, P. *J. biol. Chem.* **255**, 2540–2548 (1980).
19. Hargreaves, W. R., Giedd, K. M., Verkleij, A. & Branton, D. *J. biol. Chem.* **255**, 11965–11972 (1980).
20. Hiller, G. & Weber, K. *Nature* **266**, 181–183 (1977).
21. Goodman, S., Zagon, I. S. & Kulikowski, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7570–7574 (1981).
22. Davis, J. & Bennett, V. *J. biol. Chem.* **257**, 5816–5820 (1982).
23. Bennett, V. *Nature* **281**, 597–599 (1979).
24. Bennett, V. & Davis, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7550–7554 (1981).
25. Sattilano, R. F., Dentler, W. L. & LeCluyse, L. *J. Cell Biol.* **90**, 467–473 (1981).
26. Voter, W. A. & Erickson, H. P. *J. Cell Biol.* **91**, 331a (1981).
27. Tyler, J. M., Reinhardt, B. & Branton, D. *J. biol. Chem.* **255**, 7034 (1980).
28. Morrow, J., Speicher, D., Knowles, W., Hsu, C. & Marchesi, V. T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6592–6597 (1981).
29. Bennett, V. *J. biol. Chem.* **253**, 2292–2299 (1978).
30. Shimo-Oka, T. & Watanabe, Y. *J. Biochem.* **90**, 1297–1307 (1981).
31. Levine, J. & Willard, M. *J. Cell Biol.* **90**, 631–643 (1981).
32. Bennett, V., Davis, J. & Fowler, W. E. *Phil. Trans. R. Soc. B* (in the press).
33. Luna, E., Fowler, V., Swanson, J., Branton, D. & Taylor, D. L. *J. Cell Biol.* **88**, 396–409 (1981).
34. Mescher, M., Jose, M. & Balk, S. *Nature* **289**, 139–144 (1981).
35. Schliwa, M. & van Blerkom, J. *J. Cell Biol.* **90**, 222–235 (1981).
36. Guiley, R. L. & Reese, T. S. *J. Cell Biol.* **91**, 298–302 (1981).
37. Glenney, J., Glenney, P. & Weber, K. *Cell* **28**, 843–854 (1982).
38. Repasky, E., Granger, B. & Lazarides, E. *Cell* **29**, 821–833 (1982).
39. Burridge, K., Kelly, T. & Connel, L. *Phil. Trans. R. Soc. B* (in the press).
40. Toda, G., Oka, H., Oda, T. & Ikeda, Y. *Biochim. biophys. Acta* **413**, 52–64 (1975).
41. Fleischer, S. & Kervina, M. *Meth. Enzym.* **31A**, 15 (1974).
42. Dwyer, N. & Blobel, G. *J. Cell Biol.* **70**, 581–591 (1976).
43. Elder, J. H., Pickett, R. A., Hampton, J. & Lerner, R. A. *J. biol. Chem.* **252**, 6510–6516 (1980).
44. Fowler, W. E. & Erickson, H. P. *J. molec. Biol.* **134**, 241–249 (1979).
45. Bennett, V. & Branton, D. *J. biol. Chem.* **252**, 2753–2763 (1977).
46. Spudis, J. A. & Watt, S. J. *J. biol. Chem.* **246**, 4866–4871 (1971).
47. MacLean-Fletcher, S. & Pollard, T. D. *Biochem. biophys. Res. Commun.* **96**, 18–27 (1980).

LETTERS TO NATURE

MERLIN spectral line observations of circumstellar shells around two OH/IR stars

R. P. Norris, P. J. Diamond & R. S. Booth

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

The OH maser emission from OH/IR stars shows structure on an arc second scale, and is therefore ideally suited to observation by the new Jodrell Bank Multi Element Radio Linked Interferometer Network (MERLIN)¹ operating in the spectral line mode². Here we present MERLIN observations of two OH/IR stars: OH127.8–0.0 and OH104.9+2.4. The spatial structure of OH127.8–0.0 revealed by Booth *et al.*³ using three single base line interferometers was consistent with a simple expanding shell model. Our maps of OH127.8–0.0 confirm this interpretation, whereas the circumstellar shell that we find around OH104.9+2.4 exhibits a more complex motion, which we interpret in terms of rotation.

The sources OH127.8–0.0 and OH104.9+2.4, hereafter called OH127.8 and OH104.9 respectively, were observed at 1612.231 MHz with four telescopes of MERLIN in March 1981. The measured cross-correlation functions on the six baselines were recorded every 30 s and then Fourier transformed off-line to yield a spectrum of 80 frequency channels on each baseline. The unresolved source BL Lac was also observed to calibrate and correct the data for both amplitude and phase variations across the bandpass. During the observations on each source, additional phase variations occur which are common to all channels. These are corrected as follows.

The emission in one channel, the reference channel, containing strong emission which need not be spatially simple or unresolved was mapped using the closure phase technique⁴. The resulting spatial distribution was then Fourier transformed to obtain ϕ_s , the phase due to the source structure. The phases ϕ in the other frequency channels were then corrected using the relation

$$\phi' = \phi - \phi_{\text{ref}} + \phi_s - \phi_{\text{BP}}$$

where ϕ_{ref} is the observed phase of the reference channel and ϕ_{BP} is a small correction term describing the phase response of the instrument. The corrected phases ϕ' and the observed amplitudes were processed using the CLEAN algorithm⁵ to produce maps of the brightness distribution integrated over suitable velocity intervals.

The OH maser emission from OH/IR stars is characterized by a double-peaked spectrum, demonstrated by Fig. 1a, which

shows the autocorrelation spectrum of OH127.8. Neither OH127.8 nor OH104.9 have been detected optically, but they have been identified with the IR sources CRL230 and CRL2885 respectively. The stellar velocity is assumed to lie roughly at the midpoint of the velocities of the OH maser peaks, and the double-peaked spectrum may then be produced by expansion, contraction or rotation of circumstellar material containing the OH.

By analogy with OH/IR stars which do have optical identifications, the central star is probably an *M* giant star, but its mass may lie anywhere in the range from $\sim 1 M_{\odot}$, corresponding to a Mira variable, to $\sim 30 M_{\odot}$, corresponding to an *M* supergiant.

Figure 1a demonstrates that the flux density has decreased by approximately a factor of 3 compared with that measured in 1980³. Such variations in intensity are well known in OH/IR

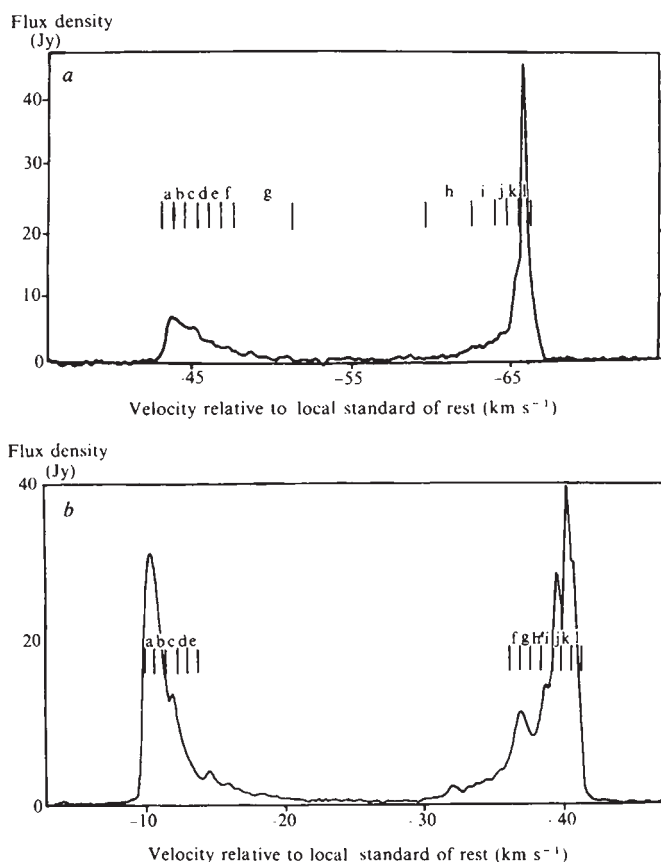
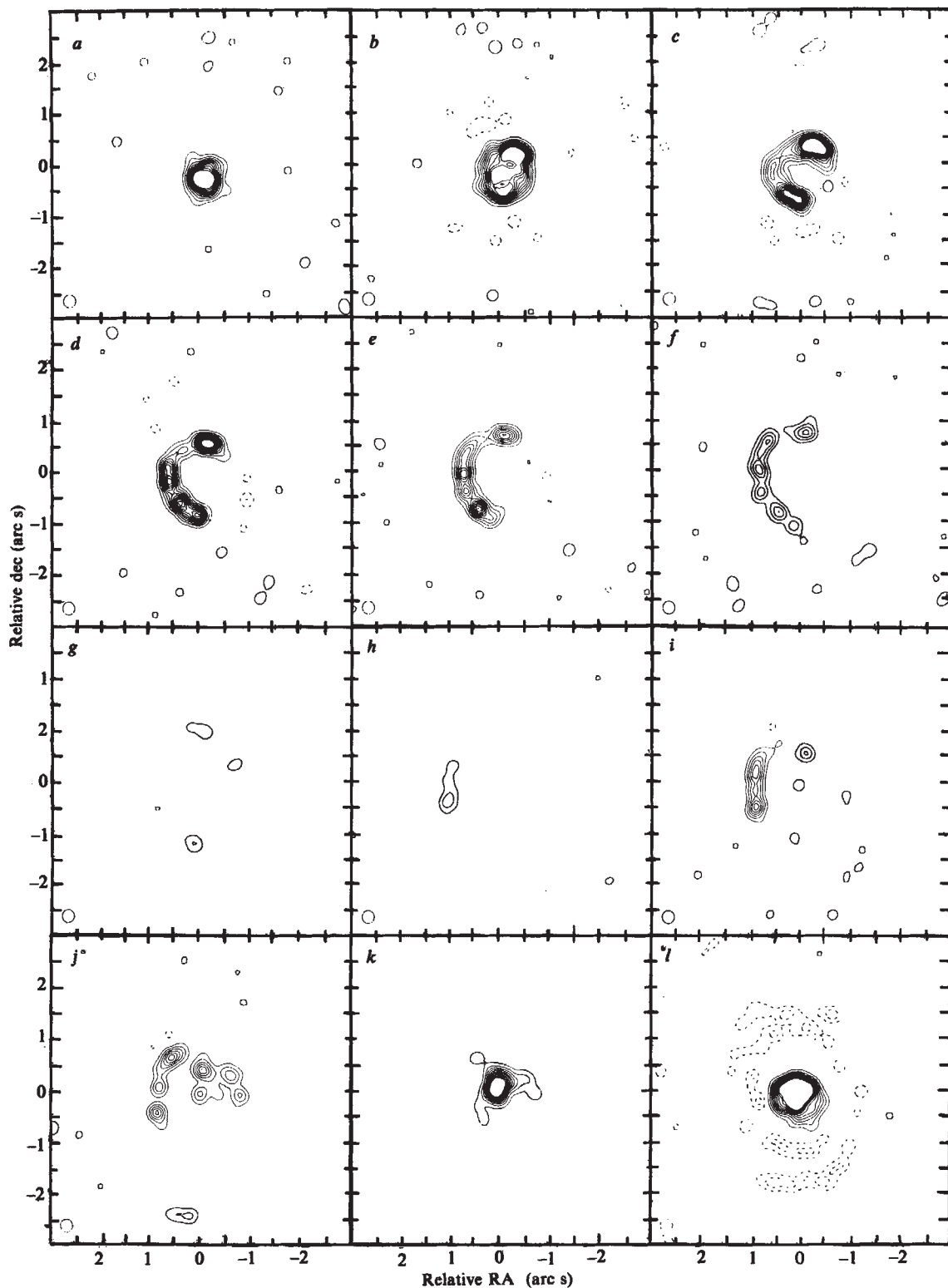


Fig. 1 Autocorrelation spectrum of: a, OH127.8; b, OH104.9 measured in July 1981.

Fig. 2 Maps of the emission in each of the 12 frequency channels shown in Fig. 1a. The lowest contour level and the contour spacing are 0.1 Jy per beam area, except for maps *k* and *l* in which the lowest contour level is 0.3 Jy. Map *l* is the reference feature mapped using the closure phase technique. Only the lowest 10 contours are shown in each case.



stars^{6,7}. Unfortunately, this reduction in brightness precludes our detecting the weaker emission detected by Booth *et al.*, preventing a detailed correspondence between the two sets of maps.

Also shown in Fig. 1a are 13 velocity ranges. Each of the maps in Fig. 2 represents the spatial distribution of the masing gas within the corresponding velocity interval in Fig. 1a. These maps have greater frequency resolution than those published by Booth *et al.*³ and confirm their earlier interpretation. The similarity of these maps to those of Booth *et al.*, despite the variation in intensity, confirms that the variations are a result of changes in the brightness of the star's IR radiation which

pumps the OH masers, rather than changes within the masing gas.

The maps show that the emission at each of the two ends of the spectrum emanates from two spatially coincident compact regions at the centre of the source, whilst the weaker emission towards the centre of the spectrum shows a roughly annular distribution. The radii of the annuli increase with decreasing velocity separation from the centre of the spectrum, in the way expected for a uniformly expanding shell of OH. The knots and gaps in the annuli indicate some turbulence or inhomogeneities in the expanding shell.

The spectrum of OH104.9, shown in Fig. 1b has a similar

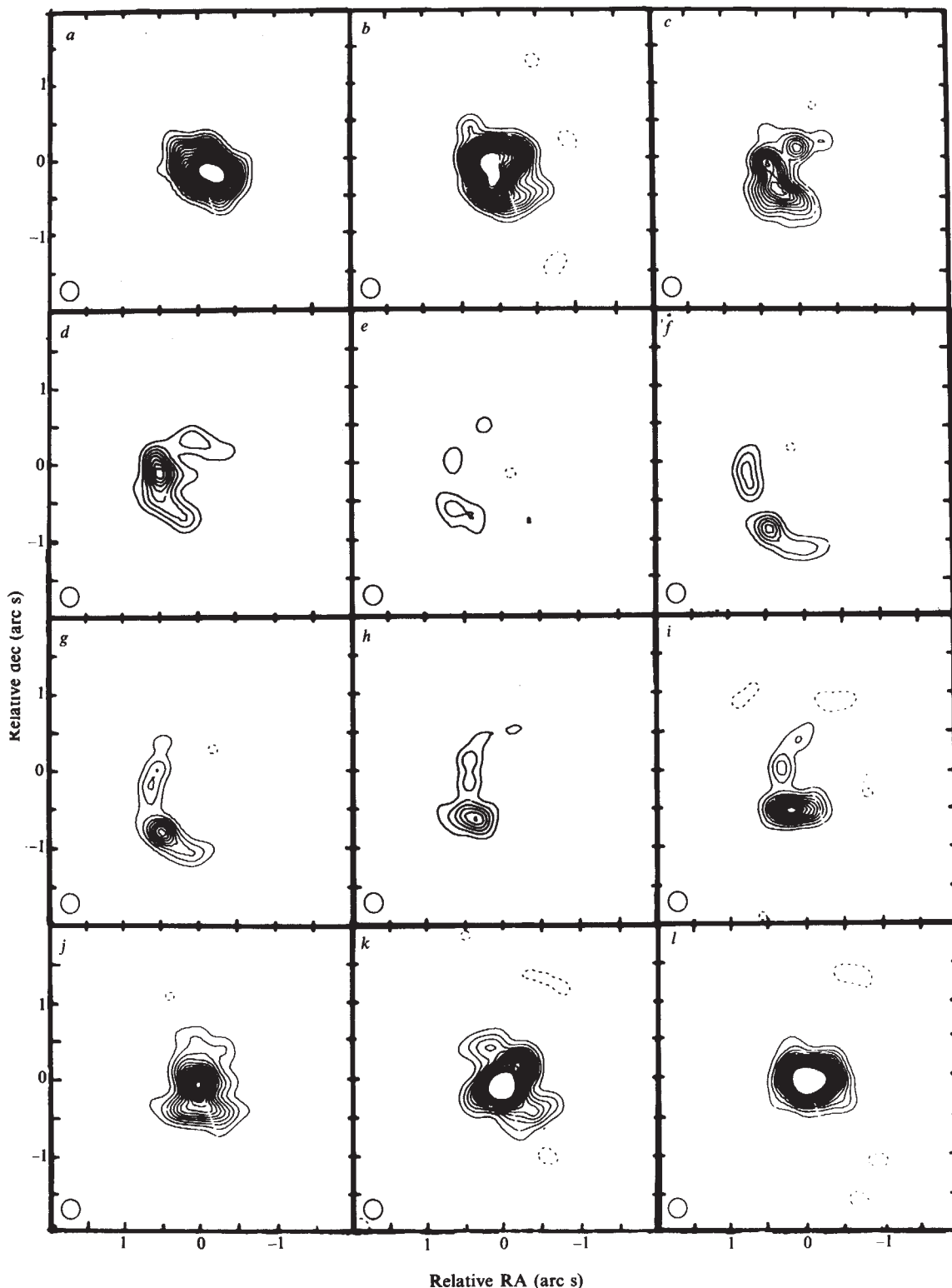


Fig. 3 Maps of the emission in each of the 12 frequency channels shown in Fig. 1b. The lowest contour level and the contour spacing are 0.5 Jy per beam area. Contours above a flux density of 10 Jy have been suppressed.

overall form to that of OH127.8 but is not as smooth and indicates a more complex motion. This is reflected in the maps shown in Fig. 3 in which no complete annuli are observed. However, the strong emission at each end of the spectrum again emanates from a region close to the centre of the source.

Note that there is no detectable emission in a region to the west of the source. However, because the maser requires a uniform line of sight velocity along its length, the maser process is extremely sensitive to inhomogeneities and anisotropies in the gas dynamics. Therefore the absence of emission may be explained as a random perturbation or non-uniformity in the gas dynamics. Such irregularities have already been seen in OH26.5 (ref. 9).

A detailed study of the components in the maps shows a further systematic departure from a simple expanding shell model. For example, in the regions to the north and south of the centre, components in the front and back of the shell which appear spatially coincident have a mean velocity of -25.8 km s^{-1} , whereas similarly coincident components to the east have a mean velocity of -25.3 km s^{-1} .

For a uniformly expanding shell, the radius $a(v)$ at which the line of sight component of the expansion velocity is v is given by

$$a(v) = R[1 - (v/v_e)^2]^{1/2}$$

where v_e is the expansion velocity and R is the shell radius.

Table 1 Parameters of the model for OH127.8 and OH104.9

	OH127.8	OH104.9
RA of centre relative to coordinate centre (arc s)	-0.11 ± 0.01	-0.12 ± 0.01
Dec. of centre relative to coordinate centre (arc s)	-0.15 ± 0.01	-0.21 ± 0.01
Position angle of rotation axis (deg)	10 ± 5	0 ± 5
Stellar velocity relative to the local standard of rest (km s^{-1})	-55.4 ± 0.2	-25.8 ± 0.2
Expansion velocity (km s^{-1})	10.9 ± 0.2	14.3 ± 0.2
Rotation velocity (km s^{-1})	0.2 ± 0.1	-0.9 ± 0.1
Shell radius (arc s)	1.53 ± 0.02	1.44 ± 0.02

These values assume that the rotation axis is perpendicular to the line of sight, because the fit is insensitive to this parameter. The quoted errors are estimated 1σ uncertainties.

For material in a uniform expanding shell, therefore, a plot of $\log a$ against $\log[1 - (v/v_e)^2]$ should yield a straight line with a gradient of 0.5. A plot of these quantities for some of the components of OH104.9 is shown in Fig. 4. The components to the north and south of the centre follow the uniform expanding shell model, but the components to the east, which are shown in Fig. 4, do not. Departure from a gradient of 0.5 may be caused either by an error in the assumed position of the expansion centre or by an incorrect value of v_e . Similarly, separation on the plot of high and low velocity components which are spatially coincident may be caused by an incorrect value of the central (stellar) velocity. However, even if all these parameters are allowed to vary, there is still no solution for a uniformly expanding, non-rotating shell which fits all the observed data.

We conclude that the kinematics of the components depart from a simple expanding shell model not only because of small-scale turbulence as mentioned above, but also because of a large scale non-radial motion. The simplest explanation of the non-radial motion is that the shell is rotating about the star as well as expanding from it. The kinematics of a rotating, expanding, thin shell are easily modelled and we have used a computer least squares program to fit such a model to the data on OH127.8 and OH104.9, resulting in the parameters given in Table 1. In each case, there is excellent agreement between the model and the data, and the r.m.s. error in velocity is $\sim 0.3 \text{ km s}^{-1}$, which is less than half our velocity resolution of 0.8 km s^{-1} . The model for OH127.8 shows no significant rotation, but the shell around OH104.9 is rotating with an equatorial speed of 0.9 km s^{-1} .

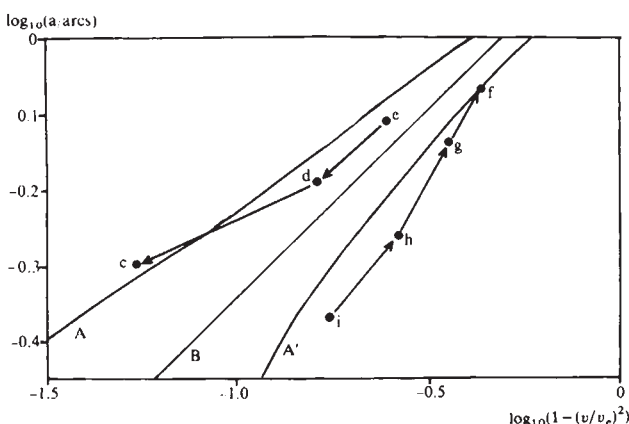


Fig. 4 A plot of $\log_{10} a$ against $\log_{10}(1 - (v/v_e)^2)$ for components at position angle 90° (see text). Line B is the curve expected from a simple expanding shell model. Lines A and A' are the curves expected from a rotating expanding shell. ●, The observed data, and the arrows connecting them are in the direction of increasing velocity. Despite the large scatter in the data points, they show clearly the bifurcation expected from a rotating expanding shell. Small letters against data points correspond to the maps in Fig. 3.

The angular momentum of the shell may be derived either from rotation of the star or from the orbital angular momentum of a binary companion. The former explanation implies a high rate of angular momentum loss from the star, leading us to favour the second explanation. A plausible model would separate the binary companion from its primary by $\sim 10^{15} \text{ cm}$, implying an orbital period of the order of hundreds of years. The current position of the companion, which may destroy the conditions for maser emission, may lie to the west of the source, where there is no detectable OH maser emission. These models are developed and discussed in more detail in ref. 10.

The evaluation of the mass loss rate from long period variable stars is of great importance in understanding stellar evolution. By assuming³ a molecular hydrogen number density of $2 \times 10^4 \text{ cm}^{-3}$, we derive a mass loss rate of $8 \times 10^{-5} M_\odot \text{ yr}^{-1}$ for OH127.8 and $5 \times 10^{-5} M_\odot \text{ yr}^{-1}$ for OH104.9. Such high mass loss rates agree well with the IR properties of OH/IR stars¹¹, although they imply a short lifetime for a star in this state.

We have produced further evidence for the evidence of a simple expanding shell around OH127.8, and have shown that OH104.9 possesses a shell which is rotating as well as expanding, and may be a member of a binary system. It is clearly necessary to extend these results to other sources, and such work is in progress.

Received 13 May; accepted 2 July 1982.

1. Davies, J. G., Anderson, B. & Morison, I. *Nature* **288**, 64–66 (1980).
2. Norris, R. P., Booth, R. S. & Diamond, P. J. *Mon. Not. R. astr. Soc.* (in the press).
3. Booth, R. S., Kus, A. J., Norris, R. P. & Porter, N. D. *Nature* **290**, 382–384 (1981).
4. Cornwell, T. J. & Wilkinson, P. N. *Mon. Not. R. astr. Soc.* **196**, 1067–1086 (1981).
5. Hogbom, J. A. *Astr. Astrophys. Suppl.* **15**, 417–426 (1974).
6. Jewell, P. R., Webber, J. C. & Snyder, L. E. *Astrophys. J. Lett.* **242**, L29–L31 (1980).
7. Herman, J. & Habing, H. in *Proc. Workshop on Physical Processes in Red Giants*, Erice (1980).
8. Norris, R. P. & Booth, R. S. *Mon. Not. R. astr. Soc.* **195**, 213–226 (1981).
9. Baud, B. *Astrophys. J. Lett.* **250**, L79–L83 (1981).
10. Diamond, P. J., Norris, R. P., Booth, R. S., Withington, S. W. & Rowland, P. R. 1982 (in preparation).
11. Werner, M. W. et al. *Astrophys. J.* **239**, 540–548 (1980).

Ion nucleation—a potential source for stratospheric aerosols

F. Arnold

Max-Planck-Institut für Kernphysik, Postfach 103 980, 6900 Heidelberg, FRG

The stratospheric aerosol layer is of considerable interest for its potential influence, at least temporarily, on the Earth's radiation budget and climate¹. While the nature of the aerosol is now reasonably well established the mechanisms leading to the formation of new particles are poorly understood^{2–4}. It is thought that the aerosols are formed by heterogeneous heteromolecular condensation involving sulphuric acid and water vapour and pre-existing condensation nuclei (CN) of tropospheric or meteoric origin. The resulting supercooled solution droplets which are composed of $\sim 75\%$ (by mass) H_2SO_4 and 25% H_2O undergo condensational and coagulative growth and are ultimately removed from the stratosphere by gravitational sedimentation. CN-formation in the stratosphere by homogeneous or ion nucleation (IN) has largely been found to be inefficient^{2–6}. I have now investigated the role of ion nucleation in stratosphere aerosol formation in the light of recent progress in both the areas of stratospheric ion composition⁷ and sulphuric acid vapour^{8–10} measurements. In contrast to previous conclusions ion nucleation is found to be a potential source for stratospheric aerosols.

Ion nucleation involves growth of ions by association of gas molecules to a critical size at which the resulting molecular cluster is stable against evaporational destruction after neutralization of its net charge by ion-ion recombination. In

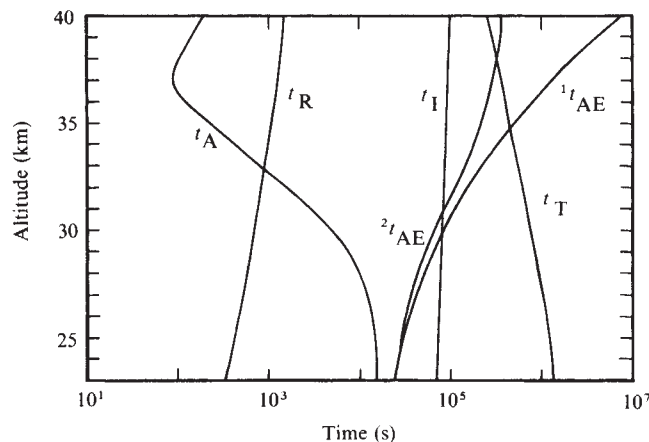


Fig. 1 Typical time scales for ion growth (t_A), ion-ion recombination (t_R), H_2SO_4 attachment to ions (t_I) and aerosols (t_{AE}) and vertical turbulent transport (t_T). For t_A two cases, H_2SO_4 - H_2O aerosols ($1t_{AE}$) and, in addition to these, meteoric smoke particles ($2t_{AE}$) (7% recondensation of meteor vapour) are given¹⁷. For details see text. For t_R evaluation measured α and $n_{\pm}$ values were taken.

stratospheric conditions only molecular clusters containing H_2SO_4 and H_2O molecules can form CN. Consequently only ions containing H_2SO_4 and H_2O molecules can be involved in IN. It has been found recently⁷ from balloon-borne composition measurements of stratospheric negative ions that the most abundant ion species above 27 km are H_2SO_4 (H_2SO_4)_n(H_2O)_m and that these already develop properties similar to H_2SO_4 - H_2O solution droplets for $n \geq 3$. The time scale for addition of an H_2SO_4 molecule to an ion is $t_A = (k[\text{H}_2\text{SO}_4])^{-1}$ where k is a phenomenological binary rate coefficient for H_2SO_4 association being typically of the order of $10^{-9} \text{ cm}^3 \text{ s}^{-1}$ in the conditions considered^{11,12} and $[\text{H}_2\text{SO}_4]$ is the concentration of H_2SO_4 -vapour molecules.

The ion lifetime compared with ion-ion recombination is $t_R = (\alpha n_{\pm})^{-1}$ where α is the effective ion-ion recombination coefficient and $n_{\pm}$ is the total charged particle concentration.

The abundance ratio $R_{n,n+1}$ for ions containing $n+1$ and n H_2SO_4 molecules is obtained from a simple steady state treatment

$$R_{n+1,n} = (1 + t_A/t_R)^{-1} \quad (1)$$

accordingly one obtains the approximation

$$R_{n_c,0} = (1 + t_A/t_R)^{-n_c} \quad (2)$$

where n_c is the number of H_2SO_4 molecules contained in an ion cluster which also remain stable against evaporational destruction after neutralization.

Thus the ion nucleation rate is

$$J = Q(1 + t_A/t_R)^{-n_c} \quad (3)$$

where Q is the galactic cosmic ray ionization rate, being of the order of $10 \text{ cm}^{-3} \text{ s}^{-1}$ in the middle stratosphere¹³.

The critical quantity n_c may be set equal to the number of H_2SO_4 molecules contained in an electrically neutral drop which is just stable and whose size may be obtained from the Kelvin expression¹⁴. The latter relates the droplet radius r_c to the supersaturation ratio $S = p/p_0$ where p and p_0 are the partial and equilibrium saturation pressures for sulphuric acid vapour. For $S \leq 6$, n_c becomes larger than ~ 10 and increases steeply as S increases. This is consistent with laboratory studies¹⁵ of IN suggesting that IN becomes efficient for $S \geq 4-6$. However, usually in these studies $t_A \ll t_R$, in contrast to the stratospheric situation.

Typical t_A and t_R values are shown in Fig. 1. Measured α and $n_{\pm}$ were taken (ref. 16, and D. Smith, personal communication) and k was set equal to $10^{-9} \text{ cm}^3 \text{ s}^{-1}$. Value for $[\text{H}_2\text{SO}_4]$ were taken from recent measurements⁸⁻¹⁰ (Fig. 2) where it was

found that $[\text{H}_2\text{SO}_4]$ below ~ 27 km is similar to the lower limit curve predicted by the theoretical model case of Turco *et al.*¹⁷ which neglects H_2SO_4 re-evaporation from the aerosol. Above ~ 35 km, $[\text{H}_2\text{SO}_4]$ is similar to the upper limit curve CHA 75) predicted by a purely photochemical diffusive model¹⁸ which neglects H_2SO_4 condensation. In the intermediate altitude range (27–35 km) $[\text{H}_2\text{SO}_4]$ is similar to $[\text{H}_2\text{SO}_4]_s$, the expected equilibrium saturation pressure for H_2SO_4 over the aerosol^{9,10}.

Thus it appears that $S < 1$ above ~ 35 km, $S = 1$ between ~ 27 and 35 km and $S > 1$ below ~ 27 km. That implies that aerosols are unstable above ~ 35 km and that IN can occur only at heights below ~ 27 km. Since $[\text{H}_2\text{SO}_4]_s$ increases strongly as temperature increases, the $[\text{H}_2\text{SO}_4]_s$ curves are different for winter and summer due to seasonal temperature changes. Corresponding $[\text{H}_2\text{SO}_4]_s$ curves for winter (45 W) and summer (45 S) and 45° latitude are also shown in Fig. 2.

Ion nucleation rates J calculated from equation (3) for average winter and summer temperatures at 45° latitude are shown in Fig. 3. Here the curve T80 (Fig. 2) was taken as a lower limit to $[\text{H}_2\text{SO}_4]$.

Generally the calculated J values increase as height increases and decrease abruptly above a certain height where S becomes smaller than ~ 10 and where consequently n_c becomes large. This is due to the large t_A/t_R ratios in this height region. Maximum J values of about $3 \times 10^{-3} \text{ cm}^{-3} \text{ s}^{-1}$ (winter) and $5 \times 10^{-4} \text{ cm}^{-3} \text{ s}^{-1}$ (summer) are reached around 29 km and 24 km respectively. Corresponding column IN rates J_c are $\sim 9.3 \times 10^2 \text{ cm}^{-2} \text{ s}^{-1}$ (winter) and $1.6 \times 10^2 \text{ cm}^{-2} \text{ s}^{-1}$ (summer). By comparison, the globally averaged J_c required to maintain the stratospheric aerosol layer is only of the order of $0.01 \text{ cm}^{-2} \text{ s}^{-1}$.

Thus IN apparently has the potential to form more CN than required to seed the stratospheric aerosol layer.

However, a large fraction of the newly formed CN may become removed before they can grow a new aerosol solution droplet. Loss processes may involve mostly CN attachment to aerosols. The lifetime of a gas molecule (molecular weight 29) against collision with aerosols t_{AE} is given in Fig. 1. Below 30 km it is of the order of only 10^4 – 10^5 s. In comparison, the time scale for adding an H_2SO_4 -gas molecule to a CN having a radius of 10^{-7} cm is of the order of 10^4 s. Thus a newly formed CN can pick up a few more H_2SO_4 -vapour molecules before it collides with an aerosol solution droplet. This is particularly true as the CN lifetime against collision with aerosols should be much larger than t_{AE} due to the large CN mass.

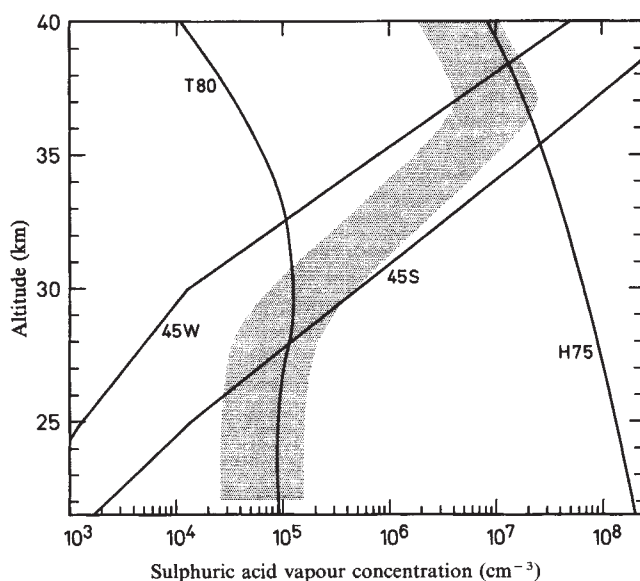


Fig. 2 Sulphuric acid vapour concentrations. Measured values⁸⁻¹⁰ are indicated by the shaded area. Extreme cases as obtained from theoretical model calculations are indicated by H75 (all sulphuric acid in the gas phase¹⁷) and T80 (no re-evaporation of sulphuric acid vapour¹⁶) for average winter and summer temperatures at 45° latitude are also shown.

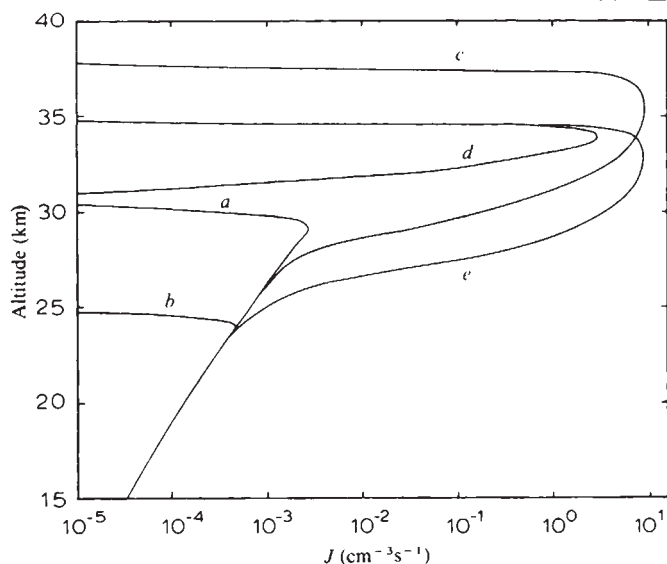


Fig. 3 Calculated ion nucleation rates J : a , average winter temperatures (T_w); b , average summer temperatures (T_s); c , cooling from T_s to T_w ; d , cooling from $T_s + 5$ to T_s ; e , cooling from $T_s + 10$ to T_s .

Assuming a CN lifetime against attachment to aerosols of 10^5 s and a yearly averaged J of the order of 10^{-3} $\text{cm}^{-3} \text{s}^{-1}$ around 25–30 km one obtains an average CN concentration $[\text{CN}]$ of the order of 100 cm^{-3} for the 25–30-km region and 10 cm^{-3} for 20 km.

Measured¹⁹ $[\text{CN}]$ are about $1\text{--}100 \text{ cm}^{-3}$ (25–30 km) and $1\text{--}10 \text{ cm}^{-3}$ (20 km) and compare reasonably well with the above model predictions for altitudes below about 25–30 km.

However, no CN formation occurs above 25–30 km in the conditions considered and therefore CN must be transported, mostly by turbulent diffusion, to this height range. The time scale for vertical turbulent diffusion t_T (per kilometre) is given in Fig. 1, being $\sim 10^5\text{--}10^6$ s (1–10 days). Thus it is much larger than the CN lifetime against collision with aerosols, which implies that little vertical CN transport should occur below ~ 35 km. Consequently, CN observed above 25–30 km can hardly be explained by the above IN model.

In the stratosphere, however, dramatic departures from a condensation–evaporation equilibrium for sulphuric acid can occur. These should be induced by short-term temperature changes which are quite common, particularly at latitudes $>45^\circ$ during local winter²⁰. An example of temperature changes at 35 km in Fig. 4 shows measured temperatures along with average summer (S), spring–autumn (SA) and winter (W) values. Typical amplitudes and time scales are about 5–10° and one to a few days respectively in winter; occasionally, dramatic temperature enhancements, ‘sudden stratospheric warmings’, occur²⁰ during which temperatures may even exceed summer values. These events are thought² to arise mostly from large amplitude planetary waves pumping heat energy from the tropical troposphere and stratosphere to the polar regions. Thus, the radiative imbalance in the atmosphere arising from the smaller amount of solar radiation absorbed at high latitudes is compensated through adiabatic processes.

As the temperature increases, $[\text{H}_2\text{SO}_4]_s$ increases (by about a factor of 10 per 10° (refs 22, 23). Eventually the entire $\text{H}_2\text{SO}_4\text{--H}_2\text{O}$ aerosol may evaporate, depending on the maximum temperature reached, in which case $[\text{H}_2\text{SO}_4]$ would be equal to curve H75 (Fig. 2). As the temperature falls again, sulphuric acid vapour may become highly supersaturated with respect to the condensed $\text{H}_2\text{SO}_4\text{--H}_2\text{O}$ phase.

The time scale for H_2SO_4 -vapour recondensation on aerosols is larger than t_{AE} (Fig. 1), being ~ 1 day around 30 km. This implies that high supersaturations may persist for at least 1 day.

A typical case for a major winter warming would be a temperature rise to average summer values and a subsequent cool-

ing to average winter values which implies a maximum S of ~ 3 (Fig. 2). Corresponding J values are shown in Fig. 3 (curve c). Around 35 km they become comparable to Q ($10 \text{ cm}^{-3} \text{s}^{-1}$) and fall off steeply above and less steeply below this height. The thickness of the IN layer is ~ 5 km. The fall-off towards larger heights is mostly due to a decrease of S causing a steep increase of n_c . The fall-off towards smaller heights is mostly due to the decrease of $[\text{H}_2\text{SO}_4]$ causing an increase of t_A/t_R .

Summer IN events for temperature increases of 5° and 10° and subsequent coolings to average summer values are also shown in Fig. 3 (curves d and e). These IN layers peak around 34 and 33 km respectively and are ~ 1.5 and 5 km thick respectively.

For curves c , d and e maximum J values are sufficiently large to allow for a rapid build-up of a limiting $[\text{CN}]$, being of the order of 10^5 cm^{-3} within <1 day. For $[\text{CN}] \geq 10^5 \text{ cm}^{-3}$ the lifetime of a CN against mutual CN coagulation becomes ≤ 1 day. For $[\text{CN}] = 10^5 \text{ cm}^{-3}$ the lifetime of an H_2SO_4 -vapour molecule against attachment to a CN is of the order of $10^4\text{--}10^5$ s. Consequently the supersaturated H_2SO_4 vapour should become consumed by condensation onto the newly formed CN within a day or less, leading to an average aerosol corresponding to only $10^3\text{--}10^4$ H_2SO_4 molecules or $\sim 4\text{--}8 \times 10^{-7}$ cm in radius.

In the days after their formation the CN undergo further mutual CN coagulation until $[\text{CN}]$ has fallen to a level of $\sim 10^3\text{--}10^4 \text{ cm}^{-3}$, at which the time scale for coagulation becomes weeks to months. At these time scales vertical turbulent diffusion becomes efficient, having a time scale t_T of the order of a few days per kilometre. Vertical turbulent diffusion leads to a broadening of the CN layer and a further decrease of the peak $[\text{CN}]$ value.

Upward diffusion carried the CN into a region where $S < 1$ leading to CN evaporation. Downward diffusion carries CN into a region of large S where they are safe from evaporational destruction by subsequent warming events.

It appears then that IN due to warming-cooling events is an efficient CN source at heights around 27–37 km. It may give rise to average $[\text{CN}]$ values of the order of $10^3\text{--}10^4 \text{ cm}^{-3}$, corresponding to average CN sizes of 1.6×10^{-6} cm and 8×10^{-7} cm respectively.

The column IN rates J_c corresponding to curves c , d and e (Fig. 3) are $\sim 4 \times 10^7 \text{ cm}^{-2} \text{s}^{-1}$, $1 \times 10^7 \text{ cm}^{-2} \text{s}^{-1}$ and $4 \times 10^7 \text{ cm}^{-2} \text{s}^{-1}$.

The column IN production for long-lived CN only (lifetime >1 day) are $\sim 6 \times 10^{10} \text{ cm}^{-2}$, $1 \times 10^{10} \text{ cm}^{-2}$ and $6 \times 10^{10} \text{ cm}^{-2}$. In comparison, the persistent IN source existing below

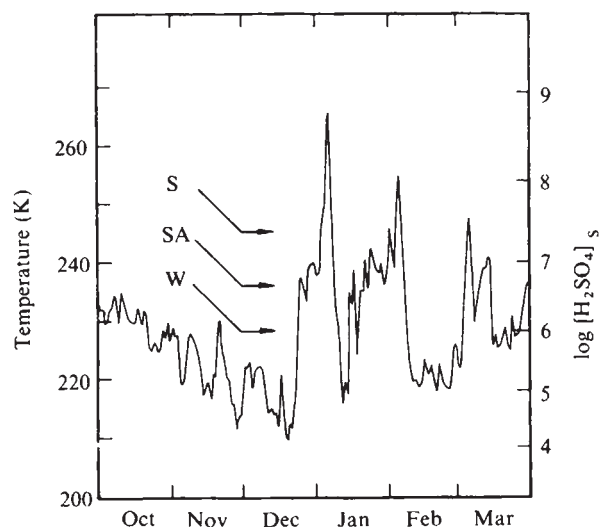


Fig. 4 Stratospheric temperatures (E. Pantzke, personal communication) as measured at 35 km altitude over Berlin (52°N , 14°E) during the winter 1975–76. Corresponding saturation concentrations for sulphuric acid vapour and average winter-, spring–autumn- and summer temperatures²⁴ are also given.

25–30 km leads to a column production of $\sim 2 \times 10^{10} \text{ cm}^{-2}$ for an entire year.

Thus the warming-cooling events contribute substantially to, or even dominate, the overall CN production. Regionally and during certain periods of time they give rise to large CN concentrations in the 27–37 km range.

Note that there is a possible influence of solar activity on the CN production by IN. Galactic cosmic radiation, the only important source of ionization in the undisturbed stratosphere, is partly shielded by the interplanetary magnetic field which gives rise to an anticorrelation of Q and solar activity. Typically Q values by about a factor of 4 over a solar sunspot cycle at middle latitudes. Since $n_{\pm} = (Q/\alpha)^{1/2}$, t_R varies by a factor of about 2. The t_R curve shown in Fig. 1 and used for the above calculations corresponds to solar sunspot maximum conditions (large t_R).

For solar sunspot minimum, t_R is smaller by a factor of 2, leading to smaller IN rates, particularly in conditions where $t_A/t_R \gg 1$. For example, the maximum value of curve a (Fig. 3) would be smaller for solar sunspot minimum conditions by about a factor of 10.

Thus the global CN production by IN may be substantially smaller during conditions of low solar activity. As a consequence the aerosol size distribution may shift towards larger sizes as fewer CN compete for the available sulphuric acid. As a change of the aerosol size distribution may affect the radiation properties of the stratospheric aerosol layer, solar activity changes may modulate the possible influence of the stratospheric aerosol layer on the Earth's radiation budget and climate.

Finally possibilities for observational tests of the ion nucleation hypothesis will be discussed. Observable quantities are [CN], $[\text{H}_2\text{SO}_4]$ and the ion composition. Concerning [CN] the IN model seems to be in harmony with measurements. In particular the CN layer recently detected by Hofmann and Rosen¹⁹ around 30–35 km is readily explained and its major features are reasonably well reproduced by the IN model.

It seems, however, that predicted [CN] values are larger than observed ones. As far as $[\text{H}_2\text{SO}_4]$ and ion composition measurements are concerned only few data exist. It would be desirable to carry out coordinated measurements of [CN], $[\text{H}_2\text{SO}_4]$ and ion composition during a major winter warming-cooling event. During the early phase of the cooling large S values and very complex ion clusters should occur while large [CN] values should persist for days after the onset of the cooling.

Note that the above model calculations of IN are only rough approximations and that extensive model calculations of stratospheric IN should be carried out.

I thank D. J. Hofmann, P. Hamill, K. Labitzke and B. Naujokat for helpful discussions, and E. Pantzke for providing stratospheric temperature data.

Received 3 March; accepted 8 June 1982.

- Pollack, J. B. *et al.* *J. geophys. Res.* **81**, 1971 (1976).
- Cadle, R. D. & Grams, G. W. *Rev. Geophys. Space Phys.* **13**, 475 (1975).
- Toon, O. B. & Farlow, N. H. A. *Rev. Earth planet. Sci.* **9**, 19 (1981).
- Castleman, A. W. Jr. *Rev. Earth planet. Sci.* **9**, 227 (1981).
- Hamill, P., Turco, R. P., Toon, O. B., Kiang, C. S. & Whitten, R. C. *J. Aerosol Sci.* (submitted).
- Mohnen, V. A. *Proc. 4th Conf. on CIAP* (US Department of Transportation, 1975).
- Arnold, F., Viggiano, A. & Schlager, H. *Nature* **297**, 371–376 (1982).
- Arnold, F. & Fabian, R. *Nature* **283**, 55 (1980).
- Arnold, F., Fabian, R. & Joos, W. *Geophys. Res. Lett.* **8**, 293 (1981).
- Arnold, F., Fabian, R., Ferguson, E. E. & Joos, W. *Planet. Space Sci.* **29**, 195 (1981).
- Viggiano, A. A., Perry, R. A., Albritton, D. L., Ferguson, E. E. & Fehsenfeld, F. C. *J. geophys. Res.* **85**, 4551 (1980).
- Webber, W. J. *Geophys. Res.* **67**, 5091 (1962).
- Thomfor, G. & Volmer, M. *Ann. Phys.* **5**, 109 (1939).
- Castleman, A. W. Jr. *Adv. Colloid Interface Sci.* **10**, 73 (1979).
- Gringel, W., Hofmann, D. J. & Rosen, J. M. *Rep. No AP-47* (University of Wyoming, Laramie, 1978).
- Turco, R. P., Toon, O. B., Hamill, P. & Whitten, R. C. *J. geophys. Res.* **86**, 113 (1981).
- Harker, A. B. *J. geophys. Res.* **80**, 3399 (1975).
- Hofmann, D. M. & Rosen, J. M. *Nature* **297**, 120 (1982).
- Labitzke, K. *J. geophys. Res.* **86**, 9665 (1981).
- Schoeberl, M. R. *Rev. Geophys. Space Phys.* **16**, 521 (1978).
- Ayers, G. F., Gillett, R. W. & Gras, J. L. *Geophys. Res. Lett.* **7**, 433 (1980).
- Rödel, W. *J. Aerosol Sci.* **10**, 375 (1979).
- U.S. Standard Atmosphere Supplements* (US Government Printing Office, Washington DC 1966).

Ozone vertical distribution and total content using a ground-based active remote sensing technique

J. Pelon & G. Megie

Service d'Aéronomie du CNRS, BP 3, 91370-Verrieres-Le-Buisson, France

The evaluation of possible perturbations of anthropogenic origin on the Earth environment has become increasingly important with the development of industrial and agricultural activities. Among these the depletion of the ozone layer under the influence of chemically active minor constituents produced by anthropogenic sources is still under question, thus emphasizing the importance of ozone monitoring in the atmosphere. Beside classical systems, such as the Dobson spectrometer and Brewer or ECC sondes, new techniques of ozone monitoring are being developed which include passive IR and microwave spectrometry and active IR and UV laser soundings¹. They should soon allow complementary routine measurements. We have previously described² the UV lidar system set up at the Observatoire de Haute Provence (44° N, 6° E) in the south of France and the methodology of the differential absorption laser technique (DIAL) which provide routine measurements of the ozone vertical distribution from the ground up to 25 km. We report here the first active ground-based measurements of the ozone distribution at higher altitudes up to 40 km and of the total ozone content using the same system.

Attempts to measure ozone number density above its maximum using a ground-based UV system are complicated because the extinction of the laser emitted light is high due to the absorption in the lower altitude layers. Furthermore the rapid decrease in ozone number density above 28 km requires very rapidly increasing acquisition times for the measurement². To overcome these major problems, a different wavelengths pair B (305.8–310.8 nm) was chosen to probe the upper levels (13–38 km) as compared with the one A (288.8–294.5 nm) previously adopted for the lower altitude range. These wavelengths choices minimize the interferences by other absorbing gases (SO_2 – NO_2) and by aerosol particles and optimize the signal-to-noise ratio². Whereas ozone distribution measurements in the boundary layer can be performed using the DIAL method³, we have restricted our measurements to the altitude range above 3 km to avoid any problem related to geometrical effects and to the large dynamic of the backscattered signals in the first few kilometres. Compared with the previously reported measurements², a factor of 5 reduction in the acquisition time has been obtained by increasing the receiving telescope diameter from 36 to 80 cm. The parameters of the lidar system are listed in Table 1. Adopted values for the ozone absorption cross-sections are those of Inn and Tanaka⁴ corrected for temperature dependence as given by Vignoux⁵.

Figure 1 represents two lidar measured ozone profiles as obtained on 1 and 2 December 1981 at 00 h UT. The acquisition time for each profile is 90 min. The vertical resolution varies between 360 m in the troposphere, 1.2 km at 25 km and 3.6 km

Table 1 Parameters of the lidar system

Emitted wavelengths		
A pair	$\lambda_1 = 288.80 \text{ nm}$	$\lambda_2 = 294.50 \text{ nm}$
B pair	$\lambda_1 = 305.80 \text{ nm}$	$\lambda_2 = 310.80 \text{ nm}$
Spectral width $\Delta\lambda < 5 \text{ pm}$		
Emitted energy $E = 40 \text{ mJ} - 10 \text{ Hz}$		
Divergence $< 0.5 \text{ mrad}$		
Telescope $\phi 80 \text{ cm}$		
Field of view 0.7 mrad		

at 35 km and above. A moving parabolic least-square regression is used to fit the experimental data². The 1σ -standard deviation is better than 5% up to 25 km and decreases down to 20% at the uppermost level. To avoid systematic errors due to the aerosol particles extinction, the aerosol layers were detected by simultaneously recording the lidar backscattered signals at 532 nm. No enhanced aerosol layer was detected during these measurements and no further correction was performed on the 0.3- μ m lidar signals². As a comparison, the ozone distribution measured using a classical Brewer-Mast sonde launched by the Météorologie Nationale at Biscarrosse (44° N, 0.5° W) on 30 November at 13 h 30 UT is also shown in Fig. 1. The agreement between the various profiles can be considered as satisfactory above 20 km, only a systematic bias <1.5% between the Brewer and lidar measurements was detected. At lower altitudes the variations observed in the ozone number density can be attributed to horizontal inhomogeneities due to dynamical motions as they are frequently observed in this altitude range: a north-south circulation was established at the 100-mbar level over western Europe in early December⁶ which could explain the variations observed. In any case, previous comparisons of lidar and ECC measured ozone profiles below 25 km have shown the reliability of the DIAL remote sensing technique⁷.

Both wavelengths pairs A and B were used to measure the ozone number density between 13 and 15 km: considering several profiles obtained during successive nights a systematic bias was detected showing ozone values 2% higher when measured using the A pair. This discrepancy can be related (1) to uncertainties in the absolute values of the cross-sections and (2) to the very high spectral resolution of the laser measurements as compared with the spectrometric ones^{4,5}. Due to experimental uncertainties, the detected bias between the A, B pairs and Brewer measurements could only be considered as estimations in this first approach; as the overall bias was not larger than 2% no final adjustment was performed.

The total ozone content can be derived from the measured vertical distribution. Various sources of uncertainty then have to be considered²:

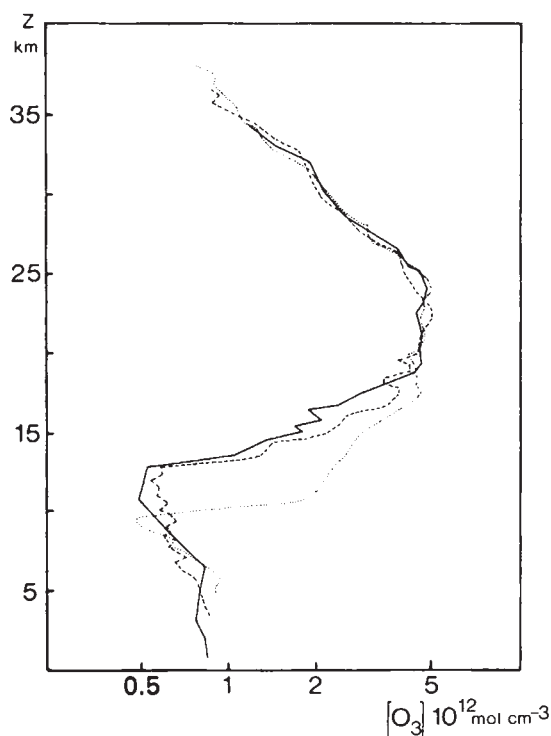


Fig. 1 Ozone concentration profiles measured by lidar at the OHP on 1 December 1981 (dashed line) and 2 (dotted line) at 0 h UT compared with the profile obtained by a Brewer sonde launched at Biscarrosse on 30 November 1981 at 13 h 30 UT (solid line).

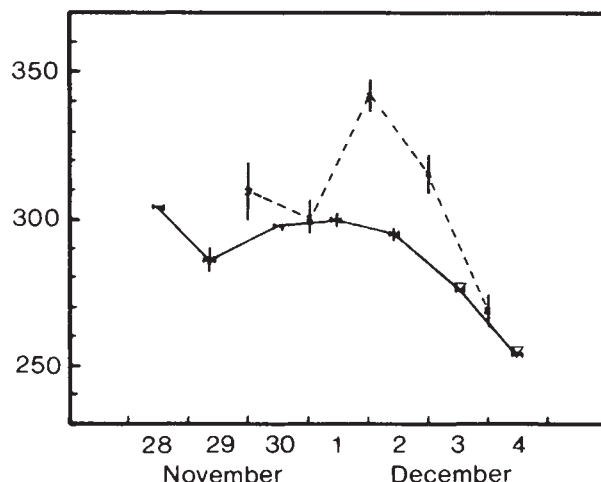


Fig. 2 Total ozone measured by lidar at the OHP (dashed line) and Dobson spectrometers at Biscarrosse (solid line) and Arosa (∇).

- (1) The statistical fluctuations of the detected signals which constitutes the major error source estimated to be < 10 Dobson units in a 95% confidence interval. This error can be further reduced by increasing the acquisition time.
- (2) The Rayleigh extinction by the atmospheric gas molecules: this uncertainty is reduced below 0.1 Dobson unit by correcting the lidar data using the pressure, temperature and humidity distribution as measured by radiosonde at the nearby meteorological station of Nîmes (44° N, 5° E).
- (3) The Mie differential extinction by aerosol particles: this uncertainty which can be large at lower altitudes² is reduced when using the A wavelength pair to measure the total content up to 15 km which leads to high ozone optical thickness as compared with the aerosol contribution. It can be estimated to be < 1 Dobson unit whereas the use of the B pair in the same altitude range will have resulted in a one order of magnitude higher uncertainty.
- (4) A systematic bias due to the discrepancies already pointed out between the two wavelength pairs and the absolute cross-sections determinations.
- (5) A constant value of 10.4 Dobson units which corresponds to the average ozone content above 40 km (ref. 8): the 2σ -standard deviation results in a 3.2 Dobson units uncertainty similar to that introduced at the lower boundary by the extrapolation of the ozone distribution down to the ground level.

The lidar derived values of the ozone total content have been plotted on Fig. 2 together with the values obtained from the Dobson spectrometer measurements performed at Arosa (46° N, 10° E) and Biscarrosse. The values obtained on 29 and 30 November and 3 December are in a good agreement with a maximum difference of 3%. On 1 December a difference of 47 Dobson units is observed which decreases down to 25 Dobson units on 2 December. It corresponds to the variations measured between 10 and 18 km when considering the vertical distributions (Fig. 1) and can thus be related as already pointed out, to the horizontal transport. Further validation of the lidar system's ability to measure the total ozone content will imply correlative measurements with a Dobson spectrometer set up at the same location which is within the scope of future developments of the geophysical station at the Observatoire de Haute Provence.

The main advantage of the remote sensing technique described above is the direct determination of the ozone number density with an independent instrument which does not require any internal or inter-calibration. This advantage can then be extended to the measurement of the total ozone content. The preliminary measurements presented here will be improved: the development of new laser sources such as excimer lasers

providing higher output energy in the same wavelength range could thus be a further determining step.

This work was supported under contract 79-71020 from the Délégation Générale de la Recherche Scientifique et Technique. The development of the O.H.P. lidar facility is supported by CNRS and INAG. We thank Professor C. Fehrenbach, Director of the Observatoire de Haute Provence for hospitality, also Dr H. Dütsch and A. Lovisa for providing the results of the Dobson spectrometers and Brewer sonde measurements.

Received 20 May; accepted 7 July 1982.

1. Global Ozone Research and Monitoring Project, Rep. 9 (WMO, Boulder, 1980).
2. Pelon, J. & Mégie, G. *J. geophys. Res.* **87**, 4947-4956 (1982).
3. Browell, E. V., Carter, A. F. & Shipley, S. T. *Proc. Quadrennial int. Ozone Symp.*, Boulder, 99-107 (1981).
4. Inn, E. C. Y. & Tanaka, Y. *J. opt. Soc. Am.* **43**, 870 (1953).
5. Vigroux, E. *Ann. Phys.* **8**, 709 (1953).
6. *European Meteorological Bulletin*, Frankfurt (1981).
7. *Upper Atmospheric Programs Bulletin*, D.E.E. (AEE 300) 81-2, 5 (Washington DC, 1981).
8. Krueger, A. J. & Minzer, R. A. *J. geophys. Res.* **81**, 4477-4481 (1976).

Copper sulphide $\text{Cu}_{1.8}\text{S}$ (Digenite I) precipitation in mild steel

J. E. Harbottle & S. B. Fisher

Central Electricity Generating Board, Berkeley Nuclear Laboratories, Berkeley, Gloucestershire GL13 9PB, UK

The microstructure of mild steel pressure vessel plate and weldments has recently been examined by transmission electron microscopy (TEM) to characterize certain numerous small ($<1,000 \text{ \AA}$) matrix precipitates and relate them to the bulk mechanical properties of the steel. The TEM examination was undertaken using in-line energy dispersive (EDAX) spectrometry. In view of the strong influence of Cu on radiation hardening and embrittlement in steels¹ particular emphasis was placed on the identification of any copper-bearing precipitates within the grains. We report here the first observation of discrete particles of copper sulphide, positively identified as $\text{Cu}_{1.8}\text{S}$ (Digenite I), and found only in the matrix of the mild steel weld specimens. In spite of having a similar copper content to the plate, the precipitation of copper sulphide uniquely in the welds is thought to arise from the high temperature oxidizing environment experienced during welding. This favours the formation of $\text{Cu}_{1.8}\text{S}$ in the presence of soluble copper from the less stable FeS commonly found in sulphur-bearing steels. The importance of this observation in relation to the mechanical properties of the steel is briefly discussed.

Examination of the microstructure of mild steel plate and weld materials, similar to those used for reactor pressure vessels, is aimed at identifying and measuring the matrix precipitate populations and relating their presence to the bulk mechanical properties. As the most numerous precipitates, and therefore those having the greatest hardening effect, are also the smallest ($<100\text{--}1,000 \text{ \AA}$) they require for identification micro-probe TEM with in-line EDAX and EELS spectrometers. These instruments permit concurrent convergent beam electron diffraction together with energy dispersive and energy-loss analyses. Electron microscope specimens in the form of extraction replicas of the steels were specially prepared and mounted to reveal precipitates having a wide compositional range. The composition typical of all the steels is: C 0.1, Mn 1.0-1.8, Si 0.1-0.4, Cr 0.05, Cu 0.05-0.2, S 0.01-0.03 wt %.

In view of the marked effect of small amounts of copper on certain post-irradiation mechanical properties^{2,3}, such as increases in yield stress and the ductile-brittle transition temperature, particular emphasis was placed on the detection and identification of copper-bearing precipitates in the grains. Special precautions were taken to eliminate stray copper signals from within the microscope. Before this study observations

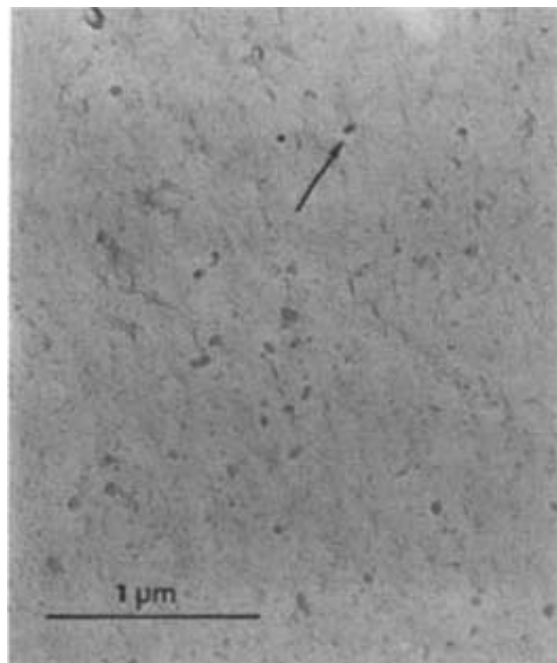


Fig. 1 Bright-field micrograph of copper sulphide $\text{Cu}_{1.8}\text{S}$ precipitation in a mild steel weld grain.

have been reported on the presence⁴, but not the crystallographic identity, of small copper-bearing precipitates in the matrix of mild steels. Small unidentified precipitates have been seen on dislocations and in the grains following ageing^{5,6} and irradiation⁷ of Fe-Cu alloys. These were assumed to be either copper precipitates or copper-vacancy complexes.

In the present study, irregularly shaped precipitates having sizes in the range $100\text{--}360 \text{ \AA}$ and a density of $\sim 10^{11} \text{ mm}^{-3}$ were observed throughout the grains of the as-welded specimens. A typical bright-field micrograph is shown in Fig. 1. These precipitates were not present in the plate material. Energy dispersive analysis indicated the presence of only copper and sulphur (Fig. 2) and the associated electron diffraction patterns (Fig. 3) consistently fitted data from the f.c.c. copper sulphide phase $\text{Cu}_{1.8}\text{S}$ (Digenite I). Copper sulphide precipitation was also found after a typical stress-relieving anneal at 600°C for 6 h, demonstrating its stability at this temperature and indicating that it was formed at high temperature during the welding process. This result is the first experimental proof of the existence of copper sulphide in steels, the crystallographic structure having been originally determined from a sample of the naturally occurring digenite mineral⁸.

The possibility that copper sulphide may form in mild steels was first suggested by Melford⁹ in a study of surface hot shortness. A feature of this effect is the localized enrichment of certain residual elements (Cu, Ni, S, Sn) that occurs below the surface scale during heating in an oxidizing environment to temperatures in excess of $\sim 1,200^\circ\text{C}$. Similar conditions will prevail during welding. A microprobe analysis of a section through the scale showed⁹ that the copper peak often coincided with an enrichment of sulphur, suggesting the presence of copper sulphide. The close association of these two elements for an individual particle near the scale was verified by mapping, although no electron diffraction pattern was obtained to provide positive identification of the particular compound.

Due to its relative insolubility ($<0.03 \text{ wt \%}$ at $1,200^\circ\text{C}$) any sulphur present in iron during steel manufacture combines to form either MnS or FeS depending on the Mn:S ratio¹⁰. Copper on the other hand is soluble up to 9 wt % in γ -iron¹¹, varying little in the temperature range $1,050\text{--}1,250^\circ\text{C}$, and therefore remains free to combine with other elements during welding, for example. Free energy calculations⁹ show that, in the

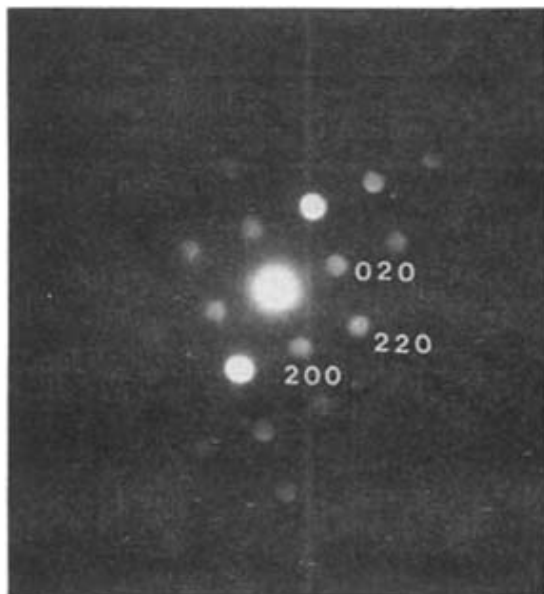


Fig. 2 Electron diffraction pattern from the arrowed particle in Fig. 1. The zone axis is [001] and the camera constant = 23.61 mm Å.

presence of oxygen and at temperatures approaching melting, copper sulphide should form from either FeS or MnS, following the reactions



and



Although we have found MnS in the large weld product inclusions but not in the weld matrix, the presence of MnO is never detected. However, FeO is commonly found throughout the matrix. This suggests that reaction (2) is the most likely to account for the observed $\text{Cu}_{1.8}\text{S}$ and FeO in the weld metal, where the high temperature, oxidizing environment would favour the transformation of FeS and the formation of copper sulphide. These conditions are not met in general in the plate material and explains the absence of both copper sulphide and wüstite in these specimens.

This result contributes principally to the understanding of copper's role in changing the yield stress of irradiated or aged ferritic steels. Previously it has been assumed that the increased hardening (or DBTT) is due to all the copper, initially in solid solution, being free to form either copper-vacancy complexes during irradiation or precipitates during thermal ageing. This

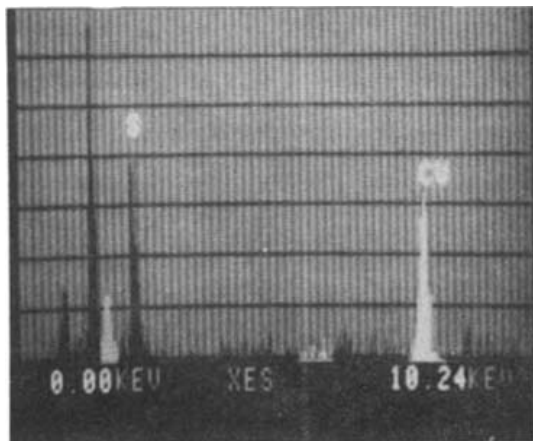


Fig. 3 EDAX spectrum from $\text{Cu}_{1.8}\text{S}$ particle. The highest peak arises from the aluminium grid holding the replica.

result shows that a proportion (up to 100% in some welds) of the total copper is precipitated as $\text{Cu}_{1.8}\text{S}$ and is therefore no longer available as free copper atoms. This can significantly reduce its influence in causing subsequent increases in yield stress. Hawthorne's¹² recent measurements support these results by showing that the addition of sulphur always reduces the detrimental effect of copper on radiation sensitivity in steels. The net change in yield stress depends on the relative quantities of free and precipitated copper. By taking the latter into account a good correlation may be obtained between the yield stress increase of various irradiated steels and (soluble copper content)^{1/2} based on a barrier hardening model.

This letter is published by permission of the Central Electricity Generating Board.

Received 9 June; accepted 13 July 1982.

1. Potapovs, U. & Hawthorne, J. R. *Nucl. Apps.* **6**, 27-46 (1969).
2. Wullaert, R. A., Shekherd, J. W. & Smith, R. W. *Am. Soc. Testing Mater. Spec. Tech. Publ.* **611**, 400 (1976).
3. Biemiller, E. C. & Byrne, S. T. *Am. Soc. Testing Mater. Spec. Tech. Publ.* **611**, 418-433 (1976).
4. Lott, R. G., Brenner, S. S., Miller, M. K. & Wolfenden, A. *Trans. Am. nucl. Soc.* **38**, 303-304 (1981).
5. Hornbogen, E. & Jung, H.-P. *Acta Metall.* **10**, 525-533 (1962).
6. Fujii, A., Nemoto, M., Suto, H. & Monma, K. *Trans. J. Inst. Metals* **9**, 374-380 (1968).
7. Takaku, H. & Tokiwa, M. *J. Nucl. Mats.* **80**, 57-68 (1979).
8. Morimoto, N. & Kullerud, G. *Am. Miner.* **48**, 110-123 (1963).
9. Melford, D. A. *J. Iron Steel Inst.* **200**, 290-299 (1962).
10. Mann, G. S. & Van Vlack, L. H. *Metall. Trans.* **B7**, 469-475 (1976).
11. Melford, D. A. *Phil. Trans. R. Soc. A295*, 89-103 (1980).
12. Hawthorne, J. R. *Trans. Am. nucl. Soc.* **38**, 304-305 (1981).

Pressure dependence of glass dissolution and nuclear waste disposal

R. A. Lewis & R. L. Segall

School of Science, Griffith University, Nathan 4111, Australia

Schemes to dispose of high-level radioactive waste by vitrification and subsequent burial call for an appraisal of the consequences of groundwater encountering the waste glass¹. Glass dissolution rates are influenced not only by properties of the glass and solvent themselves but also by properties of the combination such as the glass surface area to solvent volume ratio². Although pressure may affect glass dissolution rates, it has not received unambiguous explicit experimental investigation. A simple experiment has been designed to determine dissolution rates at room temperature of glass samples subject to very high pressure in a centrifuge. The results reported here show that the dissolution rates for Na and Si increase slightly with pressure, but that at a pressure of 160 MPa (>1,500 atm) these dissolution rates are still of the same order as those at atmospheric pressure.

Thermodynamic principles have been used to explain dissolution and redeposition under pressure in solutions at equilibrium³ but are not pertinent to dissolution kinetics which are dealt with here. Pressure effects on the solvent are not likely to be important in the dissolution double layer where large electric fields dominate⁴. Likewise, throughout this experiment, dissolution is not rapid enough for the diffusion of ions away from the surface to be a rate limiting factor. On the other hand, mass transport by diffusion in the glass might well be restricted by compression. Modification of the electronic surface states of the glass by pressure will alter their position relative to states in the solvent⁵. The application of pressure to elucidate electronic properties of semiconductors is well-known. Pressure may introduce a new, less stable phase in the glass or, in the extreme, devitrification. Hence, the study of the overall effect of pressure is germane to the development of a comprehensive theory of glass dissolution as well as having considerable practical relevance to high-level radioactive waste disposal.

Table 1 Glass composition (wt %)

SiO ₂	55.53
B ₂ O ₃	29.26
Li ₂ O	4.94
Na ₂ O	10.27
	100.00

The glass utilized was the precursor of a radioactive waste storage glass (designated 189 or M5) fabricated at AERE Harwell; its composition is given in Table 1. The stability of this simple glass in water is relatively poor. The complete high-level waste glass, which contains metallic oxides such as MgO and Al₂O₃ from Magnox waste, is considerably more durable. Beads of 2.5 mm diameter formed under surface tension by heating cubes of glass on graphite were used. The annealing sequence consisted of heating from 20 to 727 °C at 320 °C min⁻¹, holding at 727 °C for 2 min, cooling to 427 °C at 5 °C min⁻¹ and then to 20 °C at 80 °C min⁻¹. The surface is therefore well-annealed and, essentially, flame-polished. A bead was placed with 12.5 ml of deionized distilled water in each of three cellulose nitrate centrifuge tubes. The base of each tube was 152 mm from the centrifuge axis and the depth of water in each tube was 86 mm. These tubes, together with three tubes containing water alone, were spun at 20 °C in an ultracentrifuge for 45 h. The experiment was performed at centrifugation speeds 28, 35 and 40 × 10³ r.p.m. which correspond to pressures of 80, 120 and 160 MPa respectively. A control experiment consisting of six tubes was also left for 45 h without spinning. The liquid in each of the tubes was analysed for Na and Si in solution by flame atomic absorption spectrophotometry. The behaviour of the other glass constituents namely Li, an alkali and B, a glass former, would be expected to follow that of Na and Si respectively. The mean and the best estimate of the standard deviation, $\sigma(n/(n-1))^{1/2}$, for each set of data were calculated. The results for Na are presented in Fig. 1.

A small increase in the Na loss with pressure, is evident in Fig. 1. Using the Student's *t* statistic, the hypothesis that the amounts of Na measured in solution after attack at 0.1 and 160 MPa are the same is first rejected at a level of significance of 0.04. At the same level of significance the hypothesis can be rejected that the amount of Na in solution for the 160 MPa experiment compared with the one at 0.1 MPa is greater by a factor of two. It is plausible, therefore, to argue that Na leaching is slightly enhanced at the higher pressure. The analysis for Si in solution by flame atomic absorption is less accurate than that for Na, but Si exhibits a similar small increase in loss with pressure. Clearly no marked effect occurs.

Scanning electron micrographs of the glass beads before aqueous attack are basically featureless. After attack at both ambient and elevated pressures scanning electron micrographs show typical features of glass surfaces following aqueous treat-

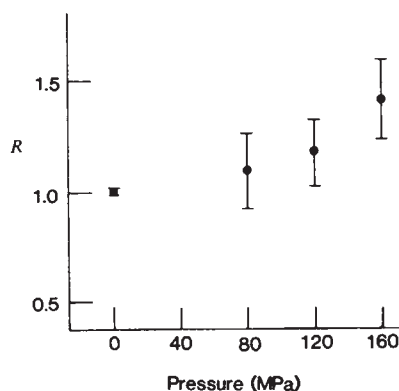


Fig. 1 The effect of pressure on glass dissolution rates, measured by the ratio, *R*, of Na loss at pressure to ambient loss.

ment—fissuring and erosion. However, there is no systematic alteration in the surface features with increasing pressure.

A general theory of the progress of glass dissolution has yet to be developed, but three stages of aqueous attack may be distinguished. For short times, the loss of material from the glass is largely the product of alkali leaching and varies as $t^{1/2}$. This regime is conveniently described by diffusion theory⁵. During intermediate times, the amount of material in solution varies as t^1 , being the result of the constant break down of the glass network. Chemical kinetics are often used to explicate the pertinent mechanisms⁶. Over long times, when the glass is either completely dissolved or the saturation of the liquid ensures redeposition at the same rate as dissolution, the amount of material in solution is constant. This state is dealt with by thermodynamics⁷.

In the present case, considering the 0.1 MPa results, the extent of Na loss is $1.42 \pm 0.03\%$ and for Si the loss is $0.7 \pm 0.3\%$. Evidently the loss of Na by exchange with H⁺ has at least occurred, and may still be occurring, when the experiment was terminated. The comparable loss of Si indicates that network break down has also been taking place. The concentrations of these ions in solution—Na, $0.59 \pm 0.01 \mu\text{g ml}^{-1}$; Si, $0.9 \pm 0.4 \mu\text{g ml}^{-1}$ —is well below saturation. In relation to the three stages of attack outlined above, in the conditions of this experiment the glass has been subject to the first and the second, but not the third stage. This may appear surprising for aqueous attack at such a low temperature (20 °C) and for such a short duration (45 h). This glass, however, is known to be very easily corroded. Other experiments in our laboratory have shown that a 10 mm × 2 mm disk of the glass dissolves entirely in <48 h at 170 °C in 15 ml of water. A variety of nuclear waste glasses show only 0.5–2% dissolution when subject to these conditions.

A difficulty in interpreting glass dissolution measurements is the large number of variables which influence the glass–solvent interaction. These include properties of the glass, such as its composition, homogeneity, melt history, environmental history, surface preparation and physical form; properties of the solvent such as its pH and species in solution; and properties of the glass–solvent system, such as the duration of contact, nature of solvent flow and temperature⁸. While it cannot be asserted that the effect of pressure will be negligible in conditions different from those of the present experiment, a knowledge of the mechanisms of the attack contributes to confidence that this is so. No conspicuous difference in the amount of Na present as a result of ambient and high pressure attack is apparent. A difference would be expected if pressure grossly affected the mechanisms occurring at the $t^{1/2}$, or diffusion stage. Likewise, the lack of marked change in the Si concentrations indicates no large effect of pressure on the mechanisms operating at the t^1 , or network break down stage.

The maximum pressure considered in this experiment is more than seven times the pressure of water at the critical point (374 °C, 22 MPa) and well above that expected in underground high-level waste repositories⁹. This experiment provides no evidence that pressure effects will be of importance in the disposal of high-level waste in glass. Absence of a pronounced pressure effect on dissolution kinetics makes a useful simplification to the complex set of variables in the glass–solvent interaction.

We thank S. Myhra, R. St. C. Smart and P. S. Turner for helpful comments. This research was supported by the Australian Institute for Nuclear Science and Engineering.

Received 17 March; accepted 6 July 1982.

- McCarthy, G. J. *et al.* *Nature* **273**, 216–217 (1978).
- Clark, D. E. & Yen-Bower, E. Luc. *Surface Sci.* **100**, 53–70 (1980).
- Durney, D. W. *Nature* **235**, 315–317 (1972).
- Morrison, S. R. in *The Chemical Physics of Surfaces*, 263–277 (Plenum, New York, 1977).
- Doremus, R. H. *J. Non-Crystalline Solids* **19**, 137–144 (1975).
- Budd, S. M. *Phys. Chem. Glasses* **2**, 111–114 (1961).
- Paul, A. J. *Mater. Sci.* **12**, 2246–2268 (1977).
- Cousens, D. R. *et al.* *Radioactive Waste Management* **2**, 143–168 (1981).
- Donath, F. A. *Nucl. Chem. Waste Management* **1**, 103–110 (1980).

Origin of sulphur and geothermometry of hydrothermal sulphides from the Galapagos Rift, 86°W

Russell Skirrow

Applied Geochemistry Research Group,
Department of Geology, Royal School of Mines, Imperial College,
London SW7 2BP, UK

Max L. Coleman

Isotope Geology Unit, Institute of Geological Sciences,
64 Gray's Inn Road, London WC1X 8NG, UK

Polymetallic hydrothermal massive sulphides from a mid-ocean ridge were first recovered in 1978^{1,2}, from the East Pacific Rise (EPR) at 21° N, and later 'black smokers' were found³ debouching fluids at 380 ± 30 °C. The Galapagos Rift and Fracture Zone Intersection (GRAFZI) Project of 1980 discovered sulphides in inactive mounds with up to 27% Cu and 31% Fe at 86° W on the Galapagos Rift⁴, that are chemically very different from the EPR 21° N sulphides (up to 50% Zn). Stable isotope studies were initiated in an attempt to find the origin of the sulphide sulphur, to determine the temperature of formation of an equilibrium assemblage of chalcopyrite (CuFeS₂) and pyrite (FeS₂), and to understand the operative seawater/rock ratio, and therefore the controls on permeability of the oceanic crust responsible for producing hydrothermal sulphides. There has been relatively little similar work on samples from 21° N (refs 5–7). We report here results which indicate sulphide formation at 390 ± 20 °C that is almost identical to that of the 21° N black smokers which were measured directly by temperature probe. The seawater sulphate contribution to the sulphur of the sulphides was much more dominant than at 21° N (ref. 7), and suggests a more fractured crust and/or more rapid hydrothermal solution flow rates, induced from higher seawater/rock ratios.

Reflected light microscopy was used to select samples in which the finely intergrown pyrite and chalcopyrite appeared to have precipitated in isotopic equilibrium. The samples were subjected to quantitative X-ray diffraction⁸ to check their purity and the relative proportions of the two minerals of interest were determined to an accuracy of ±10%. Isotopic analyses of eight samples of sulphide and two of amorphous hydrated silica were carried out in late 1981.

The two samples of amorphous silica were subjected to a citrate/dithionite-buffered leach⁹ to remove contaminant iron oxides, and then heavy mineral separation to remove intermixed sulphides, before oxygen isotope analysis. After standard preparation of sulphur dioxide¹⁰ and oxygen (by fluorination¹¹), $\delta^{34}\text{S}$ and $\delta^{18}\text{O}$ results were produced using two separate VG Micromass mass spectrometers, a 602-C double collector and a 903 triple collector, respectively. Raw data were corrected for change-over valve leakage and for isobaric interference¹² and expressed relative to the Canon Diablo Meteorite, troilite phase (CDT) for sulphur, and standard mean ocean water (SMOW) for oxygen.

The inactive hydrothermal massive sulphides were recovered from the southern margin of the northern rift valley of the double-rifted system at 86° W, and were associated with a faulted perched block⁴ (Fig. 1). The hydrothermal vent system was located at a water depth of 2,850 m and extended for at least 1 km along the base of the fault scarp. The morphological and tectonic setting of this massive Fe–Cu sulphide bears many

similarities to the ophiolite sulphide deposits of the Troodos Complex^{4,13}. Chalcopyrite-rich samples were taken from a hydrothermal veinlet at the top of the inactive sulphide mound (Fig. 2). The samples rich in pyrite were taken from the mound middle, as were the amorphous silica samples. Paragenetic relations between chalcopyrite and pyrite indicate that they precipitated in equilibrium as was suggested by Styrer *et al.*⁷ for EPR 21° N sulphides.

Mineralogy exerts a control on $\delta^{34}\text{S}$ values and there is a linear relationship for samples 4, 6, 7 and 8 of high pyrite or chalcopyrite purity (Fig. 3). The isotopic data plotted against the chalcopyrite/pyrite composition (Fig. 3) show a spread of the eight data points with values for $\delta^{34}\text{S}$ between +5.42‰ and +6.34‰. The highest isotopic values came from pyrite-rich samples from the mound middle, while the lowest values were produced by mound-top chalcopyrite-rich samples. On extrapolation to pure end-member pyrite, (100% FeS₂) a value of +6.42‰ is obtained while 100% CuFeS₂ yields +5.38‰. The isotopic fractionation factor (α) of one sulphide species with respect to the other varies as a function of temperature. The calibration of Kajiura and Krouse¹⁴,

$$1,000 \ln \alpha_{\text{Pyr}}^{\text{Pyr}} = 4.5 (10^5 T^{-2})$$

yields a temperature of formation of 388 ± 20 °C for the GRAFZI pyrite/chalcopyrite mineral assemblage. The analytical precision averages 0.04‰, but values for individual samples have been indicated by error bars in Fig. 3.

The great spread of results, with four samples deviating from linearity (Fig. 3), may be explained in terms of the mechanism of transport of subsurface hydrothermally convected seawater. In a dynamic system of convectively circulated seawater, continuously interacting with igneous rocks at different temperatures at various seawater/rock ratios, in which H₂S and SO₄²⁻ are lost from solution by precipitation of sulphide and sulphate minerals, the budget of sulphur in the hydrothermal solutions must be changing constantly. Consequently, there are two possible explanations for the isotopic composition of the samples rich in pyrite (samples 1, 2 and 5 in Fig. 3) which fall below the best-fit straight line, and are therefore depleted of ³⁴S. Biological and diagenetic activity could reset and reduce the $\delta^{34}\text{S}$. The presence of organic debris was common in both GRAFZI sulphides (data not shown) and at the EPR 21° N (ref. 2). The second mechanism would involve augmenting the hydrothermal fluids with an increased amount of basaltic sulphur ($\delta^{34}\text{S} = +0.5$ to $+1.0\%$)¹⁵ as the fluids became progressively depleted in their seawater sulphate content ($\delta^{34}\text{S} = +20\%$)¹⁵, due to precipitation of sulphide and sulphate minerals within the upper crust.

The one isotopically heavy $\delta^{34}\text{S}$ value (sample 3) is problematical. This sample may have formed as a hydrothermal solution became isotopically heavier with time, possibly due to addition of more seawater sulphate from entrained seawater or due to operation of a simple closed reservoir where residual hydrothermal solutions became enriched in ³⁴S, as the ³²S isotope preferentially left solution on sulphide precipitation. The possibility of apparent ³⁴S-enrichment in GRAFZI samples due to the presence of micro-inclusions of isotopically heavy anhydrite can be ruled out as this mineral was not found, although Arnold and Sheppard⁵ noted anhydrite in EPR sulphides.

The two samples rich in amorphous silica gave the following $\delta^{18}\text{O}$ results: sample 9, +25.2 ± 0.8‰; sample 10, +17.0 ± 0.4‰. The large errors presumably are due to the presence of water in hydrated silica which may be retained to varying extents during analysis. These results imply formation at different temperatures or mixing of a high and low temperature amorphous silica product. It is interesting to calculate the range of temperatures from these values either in equilibrium with pure seawater (SMOW) or a heavier seawater representing exchange with hot basaltic material as at 21° N for example, +1.64‰ (ref. 16). The variability in published calibrations of this equilibrium fractionation^{17,18} produce further uncertainties in the temperature calculation. If all these variations are included,

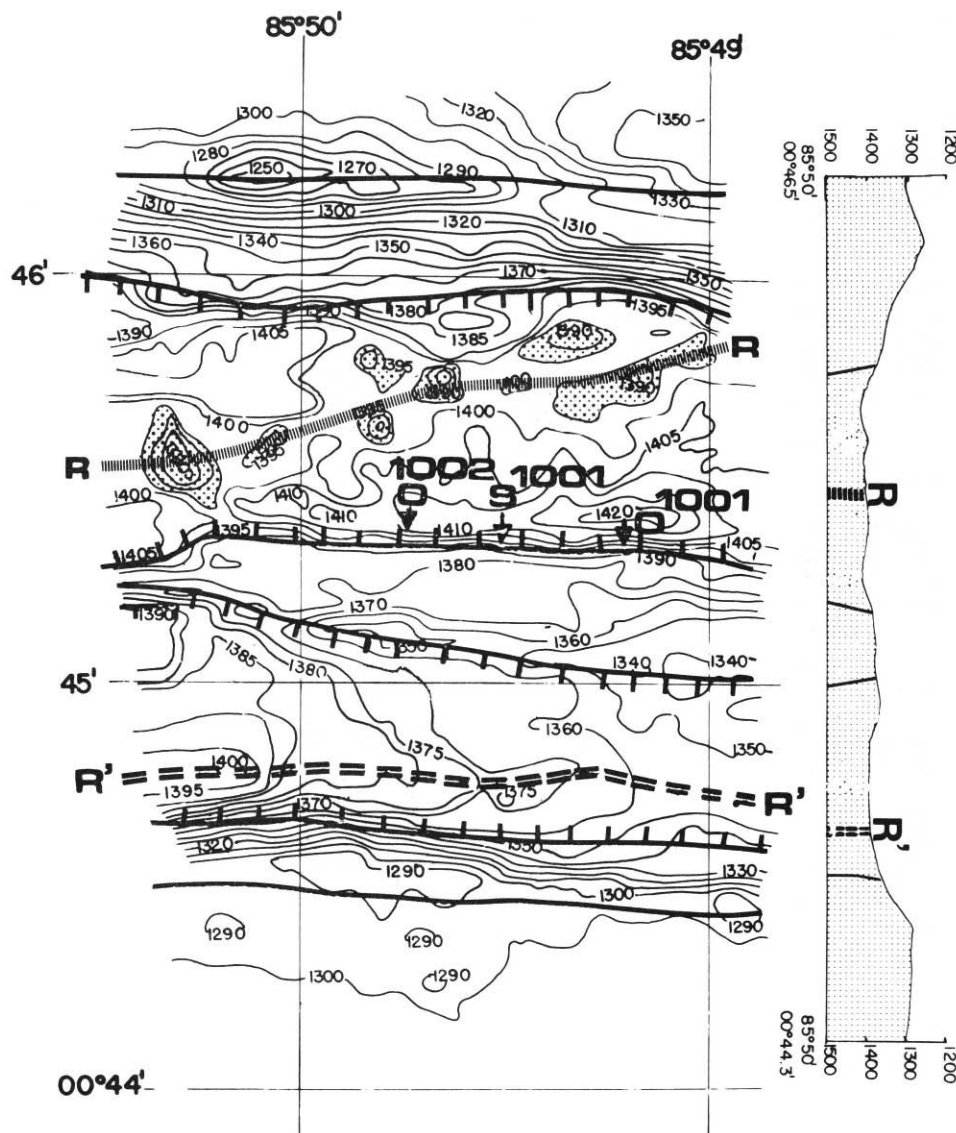


Fig. 1 High-resolution bathymetric map of the Galapagos Double Rift Area. R, principal rift axis; R', 'jumped' rift axis. The solid line delineates high crest on the tilted block and dotted areas indicate axial volcanic piles. Hashures indicate down-thrown side of normal faults; O, location of hydrothermal iron oxides; S, location of hydrothermal sulphide deposit mapped on dive 1001. (Depths in uncorrected fathoms.) After Malahoff *et al.*⁴.

the range of temperature for the two samples is from 60° to 172 °C. This suggests the possibility of considerable low temperature diagenetic additions of SiO₂ augmenting much lower temperature emissions than produced the sulphides, as has been noted for EPR 'white smokers'⁵. Amorphous silica probably formed at the periphery of the venting plume of hydrothermal water where mixing with ambient seawater occurred.

A comparison of $\delta^{34}\text{S}$ values between Recent and ancient submarine hydrothermal sulphides formed from modified seawater solutions¹⁹⁻²² reveals a similarity of composition (Fig. 4), although Kuroko deposits are associated with calc-alkaline volcanism rather than with deep-ocean tholeiitic basalts.

To compare Recent deposits in more detail, the hottest hydrothermal vents on the EPR at 21° N, at 380° ± 30 °C are actively debouching fluids which are precipitating mainly chalcopyrite, and these are very similar in temperature to GRAFZI inactive Cu-Fe sulphides which formed at 390° ± 20 °C. The overall chemistry of the two sulphide deposits is distinctively different, with the 21° N sulphides dominated by Zn, while GRAFZI deposits are richest in Cu and Fe. It is interesting to speculate whether the hottest 21° N hydrothermal vents,

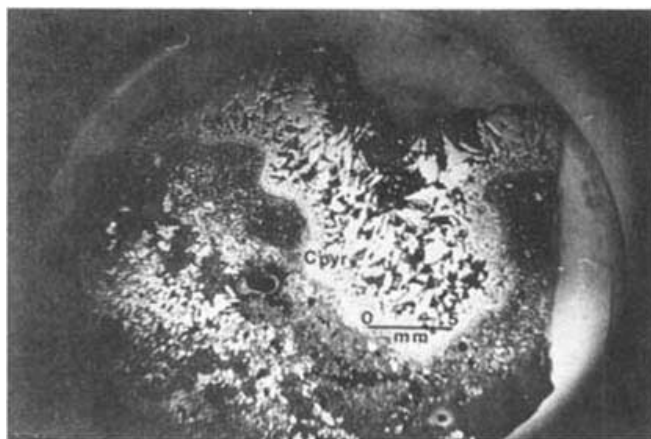


Fig. 2 Hydrothermal veinlet from the mound-top rich in 'brecciated' chalcopyrite grains (up to 3 mm). The chilled margin of the veinlet with the rapid decrease in grain size and the gossan surface are particularly noteworthy features. Scale of veinlet, 18 mm long × 12 mm wide.

previously depositing Zn sulphide, but now Cu sulphide, are sealing themselves up, as GRAFZI sulphides once did.

$\delta^{34}\text{S}$ values for 21 °N sulphides range from +1.3 to +4.1‰ and the sulphides received a 10% seawater sulphate sulphur contribution with 90% from basaltic leaching^{5,7}. Both experimental evidence²³ and now GRAFZI sulphide evidence have shown that the total hydrothermal $\delta^{34}\text{S}$ in a modern submarine deposit can exceed +5‰, and here ranges from +5.4 to +6.3‰. Styrer *et al.*⁷ observed that H_2S entirely of seawater origin may have $\delta^{34}\text{S}$ values as low as +5‰. If this is true, the entire sulphur of the GRAFZI sulphides may have a seawater sulphate origin. These sulphides are considerably enriched in ^{34}S compared with 21° N sulphides, suggesting a more dominant seawater source of sulphur at 86° W. This observation has profound implications, in that higher seawater/rock ratios than have been realised previously, can be responsible for the production of seabed hydrothermal sulphides. In terms of subsurface crustal structure, this would be manifested by a permeable, well-fractured oceanic crust and/or rapid flow rates for the hydrothermal solutions, which would allow great fluxes of seawater into the upper oceanic crust for convection and possible sea floor sulphide precipitation, on mixing with ambient bottom waters.

The analogy between GRAFZI and Cyprus-type sulphide ores can be carried further when considering stockwork hydrothermal fluid temperatures. These were found to range between 250 °C and 350 °C (ref. 24), and 300 °C (ref. 25) by different workers on Cyprus-type sulphides of the Troodos ophiolite complex. Heaton and Sheppard²⁵ postulated a seawater/rock ratio of between 1 and 5 responsible for producing disseminated stockwork cupriferous pyrites with $\delta^{34}\text{S}$, +5.1‰ (Fig. 4), while massive pyrite gave $\delta^{34}\text{S}$, +4.3‰. These values are similar to GRAFZI pyrites (up to +6.3‰) but are generally slightly lower. Spooner²⁴ suggested a major component of sulphur for the sulphides from reduced seawater sulphate at Troodos, that is,

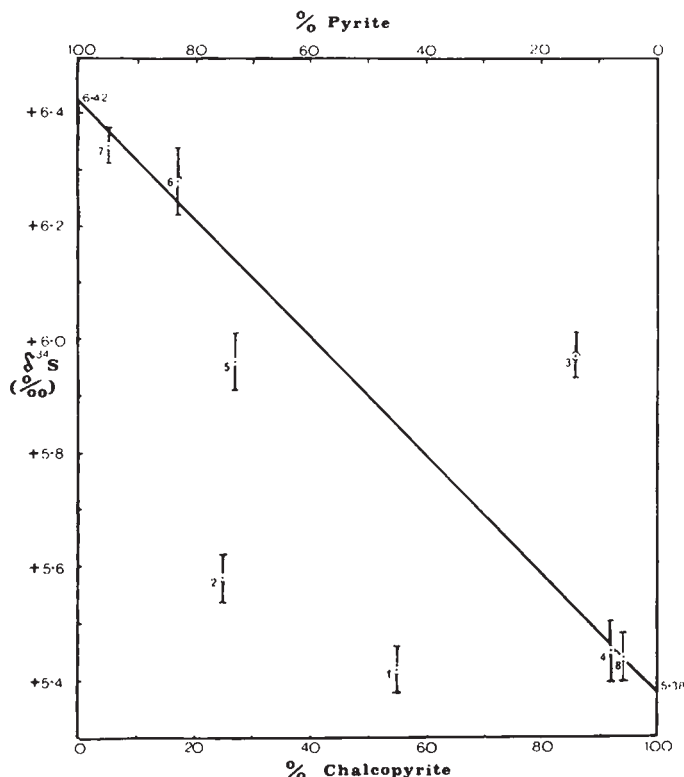


Fig. 3 $\delta^{34}\text{S}$ plotted against % chalcopyrite/pyrite mineralogy. Samples 1, 3, 4 and 8 from chalcopyrite-rich mound top. Samples 2, 5, 6 and 7 from pyrite-rich mound middle. The line is the best-fit straight line,

$$\Delta^{34}\text{S}_{\text{Pyrite-Chalcopyrite}} = \delta^{34}\text{S}_{\text{Pyrite}} - \delta^{34}\text{S}_{\text{Chalcopyrite}} = 6.42 - 5.38\text{‰} = 1.04\text{‰}$$

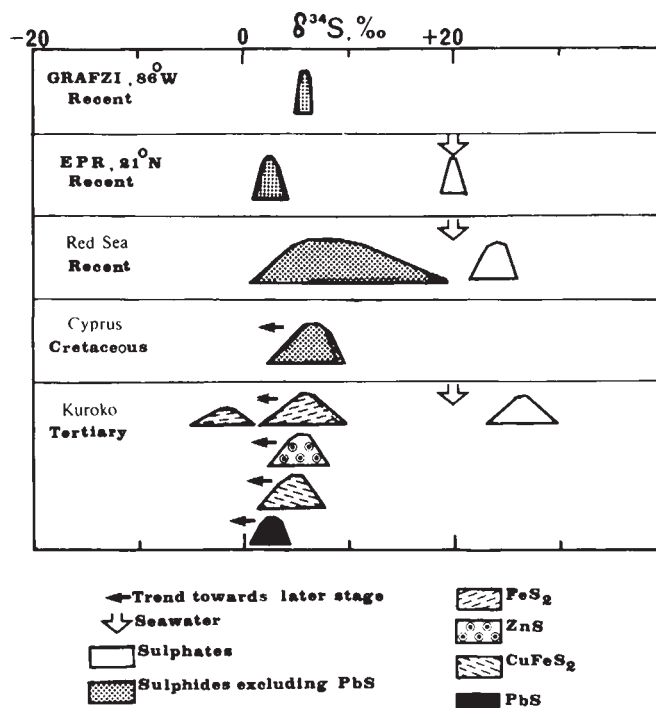


Fig. 4 Comparative sulphur isotopic data of some stratiform massive sulphide deposits associated with submarine volcanism, adapted from Ohmoto and Rye¹⁹. The sources of data on Cyprus and Kuroko are summarized in Rye and Ohmoto²⁰. The data on the Red Sea are from Kaplan *et al.*²¹, and Shanks and Bischoff²². EPR, 21° N, data are from Arnold *et al.*⁵ and Styrer *et al.*⁷.

10% and possibly up to 100%. Therefore, this seems to confirm our hypothesis of a major source of sulphur derived from seawater rather than basalt at 86° W, on the Galapagos Rift.

We thank NERC for provision of research funds and a research studentship to R.S.; Dr A. Malahoff of the National Ocean Survey; NOAA for donation of samples; Dr D. S. Cronan for invaluable discussions; Mrs M. A. Cox for analytical assistance; and Miss P. Raymond for preparation of the type-script.

Received 7 May; accepted 20 July 1982.

- Francheteau, J. *et al.* *Nature* **277**, 523–528 (1979).
- Hekinian, R., Fevrier, M., Bischoff, J. L., Picot, P. & Shanks, W. C. *Science* **207**, 1433–1444 (1980).
- Spies, F. N. *et al.* *Science* **207**, 1421–1433 (1980).
- Malahoff, A., Cronan, D. S., Skirrow, R. & Embley, R. W. *Mar. Mining* (in the press).
- Arnold, M. & Sheppard, S. M. F. *Earth planet. Sci. Lett.* **56**, 148–156 (1981).
- Haymon, R. M. & Kastner, M. *Earth planet. Sci. Lett.* **53**, 363–381 (1981).
- Styrer, M. M. *et al.* *Earth planet. Sci. Lett.* **53**, 382–390 (1981).
- Tatlock, D. B. *U.S. geol. Surv. Bull.*, 1209 (1966).
- Jackson, M. L. *Soil Chemical Analysis* (Constable, London, 1958).
- Robinson, B. W. & Kusakabe, M. *Analyt. Chem.* **47**, 1179–1181 (1975).
- Clayton, R. N. & Mayeda, T. K. *Geochim. cosmochim. Acta* **27**, 43–52 (1963).
- Coleman, M. L. *IGS Stable Isotope Rep. No. 45* (1980).
- Adamides, N. G. *Proc. Int. Ophiolite Symp., Geol. Surv., Cyprus*, 117–127 (1979).
- Kajiura, Y. & Krouse, H. R. *Can. J. Earth. Sci.* **8**, 1397–1408 (1971).
- Sakai, H., Veda, A., & Field, G. W. *U.S. Geol. Surv. Open-File Rep.* 78–101 (1978).
- Craig, H. *et al.* *EOS* **61**, 392 (1980).
- Bottinga, J. & Javoy, M. *Earth planet. Sci. Lett.* **20**, 250–265 (1973).
- Matsuhisa, Y., Goldsmith, J. R., & Clayton, R. N. *Geochim. cosmochim. Acta* **43**, 1131–1140 (1979).
- Ohmoto, H. & Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits*, 2nd edn (ed. Barnes, H. L.) 509–567 (Wiley, New York, 1979).
- Rye, R. O. & Ohmoto, H. *Econ. Geol.* **69**, 826–842 (1974).
- Kaplan, I. R., Sweeney, R. E. & Nissenbaum, A. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, T. & Ross, D. A.) 474–498 (Springer, New York, 1969).
- Shanks, W. C. & Bischoff, J. L. *Geol. Soc. Am. Ann. Mtg. Abstr.* 1266 (1975).
- Ohmoto, H., Cole, D. R. & Mottl, M. J. *EOS* **57**, 342 (1976).
- Spooner, E. T. C. in *Volcanic Processes in Ore Genesis* 58–71 (Institute of Mining and Metallurgy, London, 1976).
- Heaton, T. H. E. & Sheppard, S. M. F. in *Volcanic Processes in Ore Genesis*, 58–71 (Institute of Mining and Metallurgy, London, 1976).

Fluxes of biogenic components from sediment trap deployment in circumpolar waters of the Drake Passage

Gerold Wefer*, Erwin Suess†, Wolfgang Balzer‡, Gerd Liebezeit‡, Peter J. Müller*, C. André Ungerer† & Walter Zenk‡

* Geologisches Institut, Olshausenstrasse 40-60, 2300 Kiel, FRG

† School of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

‡ Institut für Meereskunde, Düsternbrooker Weg 20, 2300 Kiel, FRG

Circumpolar surface waters dominate the circulation of the Southern Ocean and sustain one of the ocean's largest standing stocks of biomass thereby producing a significant output of biogenic components, mainly diatoms, to the bottom sediments. Generally transit of biogenic matter from the sea surface to the sea floor affects nutrient regeneration fuels benthic life and transfers signals to the sediment record¹⁻⁵. Reliable quantification of the relationship between biological production, fractionation of skeletal and tissue components and bottom sediment accumulation depends on direct vertical flux measurements from sediment trap deployments⁶⁻⁹, which have proved to be most scientifically productive¹⁰⁻¹³. We now present data on vertical mass fluxes from the Southern Ocean and evidence for strong biogeochemical fractionation between organic carbon-, nitrogen- and phosphorus-containing compounds, siliceous and calcareous skeletal remains, and refractory aluminosilicates.

Results are from the first deployment of sediment traps in Antarctic waters. Our traps, with a collection area of 314 cm² (ref. 14), were attached to a moored array located at 60°54.6' S and 57°06.0' W in 3,625 m of water depth for 52 days from 2 December 1980 to 25 January 1981. During that time the integrated mean primary production in the study area was 0.1 g C m⁻² day⁻¹ (ref. 15). Current and temperature observations at 400 m, 1,500 m and 3,590 m of water depth revealed some unexpected results. The upper two current meters showed a general tendency of eastward water movement, although the short time series yielded no reasonable mean current vectors. The mean temperatures recorded were 1.76 °C and 1.20 °C at 400 m and 1,500 m, respectively. In both cases the standard deviation was 0.09 °C from 5,165 individual values.

In contrast to these mid-water current observations, which characterize the southern extension of the Antarctic Circumpolar Current in the Drake Passage, an unexpectedly persistent undercurrent was measured at abyssal depth. This movement feeds near-bottom water from the Scotia Sea continuously into the deep South Pacific Ocean, has a mean vector of 20.9 ± 5.3 cm s⁻¹ towards 222° relative to true north and maximum velocities of 32 cm s⁻¹ with an average temperature of -0.02 ± 0.06 °C. The countercurrent was also noted north of Livingston Island during the Dynamic Response and Kinematic Experiment (DRAKE) in 1979¹⁶.

Material in the sediment traps was recovered from water depths at 965 m and 2,540 m; the samples are designated M 269-965 and M 269-2540, respectively. Chloroform for preservation was released into the collecting vessel during trap deployment from an imploding glass vial. Surface particulate matter was also collected at the mooring site on 25 January 1981 for 8 h. About 3 m³ of sea-surface waters from the RV *Meteor's* contamination-free pumping system were continuously filtered through a 75-µm mesh sieve to concentrate enough material for individual component analyses; this sample is designated M 8101-38.

On recovery the sediment trap material and the surface particulate matter were split into 10 aliquots by pipetting from a suspension in filtered seawater which was kept homogenous

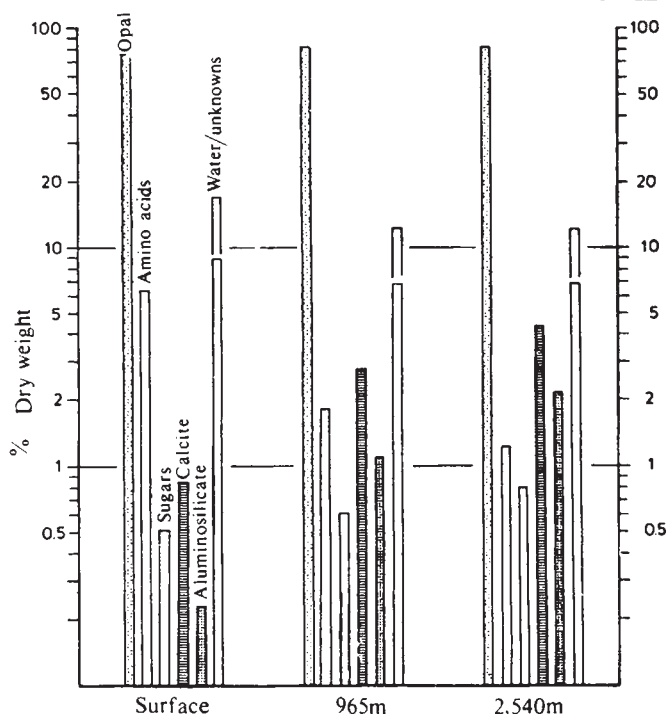


Fig. 1 Relative changes in major phase composition of particulate matter with water depth in the Drake Passage.

by continuous agitation. To minimize bacterial degradation, splits were immediately poisoned with 100 p.p.m. HgCl₂ and frozen. Just before analysis, the samples were thawed and de-salted by repeated washings in distilled water, then they were freeze-dried. The wash water was retained for labile phosphorous determinations. The yields of 8 out of 10 splits gave total dry weights, estimated to be accurate to within 5 mg. These estimates were then used to calculate the total mass fluxes of 143 ± 1 g m⁻² yr⁻¹ at 2,540 m, 163 ± 2 g m⁻² yr⁻¹ at 965 m and 471 ± 6 g m⁻² yr⁻¹ at the surface. The measure for the mass flux at the surface is the product of the mean organic carbon fixation rate¹⁵ multiplied by the bulk mass of the surface particulate matter divided by the concentration of organic carbon, and is considered a first approximation of the net bulk production. All individual component fluxes will be calculated from these bulk flux estimates.

Individual splits were analysed for elemental and phase composition, amino acid, amino sugar and carbohydrate contents and foraminiferal counts, using methods described elsewhere¹⁷. In addition, hand-picked foraminifera were analysed for their stable isotope composition as described elsewhere¹⁸. Results of the major element and stable isotope compositions are listed in Table 1.

Most of the material collected in the traps, and virtually all of the surface particulate matter, was biogenic with a very minor portion of aluminosilicates. Figure 1 illustrates the relative changes in the major particulate phases with water depth. Whole individuals of diatoms, fragments and molds of crustaceans and tests of foraminifera were most abundant. The tests were almost exclusively of the left-coiling variety of the species *Globoquadrina pachyderma*. Faecal pellets were also common.

Quantitatively, diatomaceous debris is the dominant constituent as seen in the high SiO₂-content (Table 1, Fig. 1). Hereby the percentages listed refer to anhydrous SiO₂ only, thus the H₂O of diatomaceous opal, estimated to be between 6 and 9 wt%, may largely account for the unknown portion of the total biogenic matter.

The next most important phase is organic matter which shows a change from the surface particulates to the deep trap material in decreasing protein and labile phosphorus and increasing carbohydrate contents. Calcium carbonate and the aluminosilicate fractions increase in relative importance with water depth, obviously due to the more rapid disappearance of opal and

Table 1 Major element and stable isotope composition of sediment trap material and surface plankton

(wt%) de-salted and freeze-dried material									$\delta^{18}\text{O}$	$\delta^{13}\text{C}$
	C-total C-carb C-org	N-total	P-sum P-leach P-HCl P-total	Ash SiO ₂	Ca-sol Ca-res	Mg-sol Mg-res	Al-sol Al-res	Ba-sol Ba-res	(‰ PDB) >300* ~200 ~125	
M 8101-38, sea surface										
Split (1)	8.00 ND ND	1.35 1.39	0.655 0.578 0.015 0.077	78.0 74.9	0.194 0.014	0.155 0.024	0.010 0.023	0.002 0.007	ND	ND
Split (8)	7.78 0.14 7.12 [†]	1.34 1.32	0.589 0.513 0.014 0.076	78.5 75.5	0.197 0.016	0.140 0.022	0.005 0.021	0.002 0.003	ND	ND
M 269-965, 965 m depth										
Split (1)	3.36 ND ND	0.48 ND	0.064 0.024 0.005 0.041	86.4 81.5	0.832 0.095	0.067 0.098	0.005 0.106	0.023 0.094	3.26 3.08 3.39	0.74 0.83 0.80
Split (8)	3.85 0.33 2.85 [†]	0.51 ND	0.066 0.026 0.007 0.041	86.2 80.0	0.737 0.058	0.072 0.060	0.021 0.122	0.021 0.124	ND	ND
M 269-2540, 2540 m depth										
Split (1)	4.19 ND ND	0.48 ND	0.090 0.046 0.008 0.044	85.8 79.7	1.18 0.152	0.088 0.110	0.041 0.198	0.033 0.045	3.28 3.44 3.04	1.06 0.67 0.76
Split (8)	3.49 0.55 2.76 [†]	0.42 ND	0.102 0.064 0.010 0.038	86.7 78.4	1.38 0.167	0.096 0.122	0.048 0.223	0.034 0.076	ND	ND

In each column the sequence of numbers for one sample split refers to the sequence of components as listed in the column heading: N-total, duplicate Kjeldahl determinations; C-total, CO₂ liberated by heat combustion; C-carb, CO₂ liberated by phosphoric acid; C-org, CO₂ liberated by dichromate/sulphuric acid digestion; P-HCl, hydrochloric acid-soluble; P-leach, water soluble; P-total, after water leach by nitric/fluoric acid digestion; P-sum, water soluble plus total phosphorus (calculated); ash, after ignition at 400 °C for 2 h; sol, hydrochloric acid soluble; res, residual (not hydrochloric acid soluble); for analytical scheme see ref. 17. ND, not determined.

* Splits (3)(6); shell size in μm .

† C-org contents used in carbon budget of Table 2.

organic matter. This is evident when comparing the respective changes in component flux rates with water depth.

The total calcium carbonate flux (determined from Ca and CO₂ analyses) into the traps was 4.5 and 6.6 g m⁻² yr⁻¹, respectively; whereas the *G. pachyderma* test flux (determined from test counts) was $\sim 10 \times 10^5$ individuals for each trap. For the test sizes 125, 200 and 300 μm , individual weights ranged from 1.1 to 8.9 μg with a mean test weight of 5.2 μg , thus when comparing both fluxes, it appears that *G. pachyderma* may account for the entire calcium carbonate flux.

The stable isotope values of *G. pachyderma* for three size fractions are also listed in Table 1. The tests in both traps showed mean $\delta^{18}\text{O}$ values of 3.25‰ PDB and appear unchanged during growth. This is in contrast to results from other neritic environments where, in general, for the same species the larger (older) shells are enriched in ^{18}O (ref. 19). The constant $\delta^{18}\text{O}$ values may be due to the rather small temperature variation within the euphotic zone of the study area. Depth migration during growth, if it occurs in the life cycle of *G. pachyderma*, does not lead the organism to encounter large temperature changes in circumpolar waters. The mean $\delta^{18}\text{O}$ value of 3.25‰ PDB indicates an equilibrium precipitation temperature of +1.8 °C based on the 'calcite' palaeotemperature equation of Epstein²⁰ and assuming -0.65‰ for the circumpolar surface water²¹. The calculated temperature is in good agreement with measured ones which we have recorded; that is +3.0 °C at the sea surface, -0.26 °C at 100 m, +1.44 °C at 203 m and +2.01 °C at 446 m of water depth (R. Wittstock, personal communication). The salinity varies over this same depth range between

Table 2 Organic carbon, nitrogen and phosphorus budgets of surface plankton and sediment trap material

	Sea surface M 8101-38	965 m depth M 269-965	2,540 m depth M 269-2540
	($\mu\text{mol g}^{-1}$ dry wt)		
Organic carbon	5,930	2,380	2,300
Amino acid carbon	2,895	713	531
Amino sugar carbon	46	42	40
Carbohydrate carbon	170	207	256
Carbon accounted for	52%	40%	36%
Organic nitrogen	964	357	321
Amino acid nitrogen	748	193	145
Amino sugar nitrogen	7.7	7.1	6.6
Ammonia nitrogen	83	33	25
Nitrogen accounted for	87%	65%	55%
Total phosphorus*	201	21	31
Water-soluble	176	8	18
HCl-soluble*	5	2	3
Insoluble	20	11	10
Phosphorus accounted for	100%	100%	100%

Amino acid carbon and nitrogen, sum of 22 individual amino acids; amino sugars, sum of galactosamine and glucosamine; carbohydrate carbon, sum of 11 hexoses; results of individual amino acid and sugar spectra will be published elsewhere (P.J.M. and G.L. in preparation).

* Organic+apatitic phosphorus, whereby HCl-soluble fraction may be considered apatitic²⁴.

Table 3 Mean annual production and flux rate estimates of individual biogenic components and aluminosilicates at the sea surface, at 965 m and 2,540 m of water depth at station M 269

Particulates	Sea surface M 8101-38	965 m depth M 269-965	2,540 m depth M 269-2540
		(g m ⁻² yr ⁻¹)	
Total mass flux	471	163	143
Organic carbon	36.5	5.41	4.78
Amino acid carbon	16.4	1.39	0.91
Carbohydrate carbon	0.96	0.40	0.44
Amino sugar carbon	0.26	0.08	0.07
Organic nitrogen	6.36	0.81	0.64
Amino acid nitrogen	4.93	0.44	0.29
Ammonia nitrogen	0.55	0.075	0.050
Amino sugar nitrogen	0.05	0.016	0.013
Organic phosphorus			
Water soluble	2.57	0.041	0.079
Insoluble	0.29	0.057	0.046
Silica, anhydrous	354	132	113
Calcium carbonate	5.5	4.5	6.6
Barite	0.06	0.36	0.23
Aluminosilicates	1.08	1.79	3.00
Al ₂ O ₃	0.20	0.35	0.57
MgO	0.18	0.21	0.27
CaO	0.10	0.17	0.32

Organic phosphorus (insoluble), P-total minus P-hydrochloric acid soluble; calcium carbonate, mixed Ca-Mg-carbonate probably also containing SrCO₃; barite, soluble plus residual barium plus stoichiometric SO₄; aluminosilicates, sum of residual oxides plus 3 × (Al₂O₃) as aluminosilicate SiO₂, flux rates do not include quartz-SiO₂; the individual component fluxes are calculated by averaging the compositions.

33.7 and 33.5% S, therefore our data indicate that *G. pachyderma* precipitates its tests close to isotopic equilibrium. The carbon isotope values also do not vary with shell size and thus do not show the 'normal' trend of ¹³C enrichment with increasing age¹⁹, again the rather homogeneous water column properties seem responsible for this low variability in the stable carbon isotope signature of *G. pachyderma*.

Dramatic qualitative and quantitative changes with water depth were observed for the organic matter fraction (Table 2). Whereas in surface plankton 52% of the total organic carbon may be accounted for as being tied up in polypeptides, proteins and polysaccharides, these portions decrease with depth to 40% and 36%, respectively. The non-characterized fraction probably contains hydrocarbons, wax esters and other lipid material as well as 'humic matter'; its proportion increases with depth from ~50 to ~65% of the total organic carbon pool.

Because only < 1 µmol g⁻¹ of ammonia-N could be contained in the aluminosilicate fraction as fixed NH₃ (ref. 22), we assume that all nitrogen is organically bound. The free ammonia-N listed in Table 2 is also organically bound but is liberated from the protein fraction during hydrolysis.

The decrease in amino yield and the increase in carbohydrate contents with water depth reflect the preferential degradation of proteinaceous material²³. This is supported by a decrease in both amino acid-C and amino acid-N, whereas the relative increase in the non-characterized organic nitrogen fraction might reflect a pronounced abundance of nucleic acid bases and/or 'humic material'. Amino sugars on the other hand, comprise a relatively stable fraction of both organic-C and organic-N at all depths, indicating that they may be derived from poorly degradable chitinous matter.

For phosphorus we attempted partitioning by a methodology based on fractional solubilization. This approach explains why we account for the entire phosphorus budget. Hereby the water-soluble portion probably represents very labile organic phosphorus compounds such as ATP and related energy storage compounds and nucleic acids, released from living cells. We then defined a small fraction of phosphorus, soluble in hydro-

chloric acid, which could come from apatitic biogenous skeletons²⁴. Consequently, the calculated difference between P-sum minus P-HCl and P-leach represents an insoluble organic phosphorus fraction such as phospholipids from cell walls or similar tissue structures. The relative increase of this fraction from 10% at the sea surface to ~40% at depth also indicates it to be of a more refractory nature.

We may now compare flux rate changes of organic phases with mineral phases during transit through the water column (Table 3), because ultimately, fractionation resulting from such differential flux rates transforms surface particulate matter into bottom sediments. From Table 3, four types of biogeochemical regeneration can be differentiated. (1) Labile organic phosphorus is the most rapidly recycled; only < 5% of that fraction escapes recycling within near-surface waters and reaches depths of ~1,000 m, below which its flux remains essentially constant. (2) Polypeptides, presumably proteinaceous material of surface plankton are at first rapidly recycled between the surface and ~1,000 m of depth and thereafter at a slower rate. This exponential degradation is seen in changes of amino acid-C, amino acid-N and ammonia-N fluxes; it essentially affects the entire particulate nitrogen pool. (3) Fluxes of skeletal silica, polysaccharides, amino sugars and insoluble organic phosphorus show a slight tendency of refractory chemical behaviour which is marked by imperceptibly small flux rate changes in deep water. This is not surprising, as all are 'residual' components of biogenic matter: opal of diatom tests, carbohydrates and insoluble phosphorus of cell walls and amino sugars of chitinous skeletons. (4) Finally, the behaviour of calcium carbonate, barite and aluminosilicates cannot be adequately assessed from our flux data, because their sources are not restricted to surface waters of the same region as indicated by the deep countercurrent. Therefore, the bulk production rate assumed for the surface is a poor estimate. All three components of this group may have sources farther to the east of the sampling site and thus are not included in our production rate estimate from surface particulate matter. Similarly, aluminosilicates are advected horizontally and vertically throughout the water column, and the deeper the traps are deployed, the larger the aluminosilicate fluxes will be. But at least our flux data indicate that calcium carbonate, barite and aluminosilicates are very refractory and will undergo little, if any, recycling in the water column below ~1,000 m due to rapid transit.

This work was supported by the Deutsche Forschungsgemeinschaft and the US Office of Naval Research through grant N00014-79-C-0004, Project NR083-1026 to Oregon State University. We thank especially B. von Bodungen and R. Wittstock for help in deployment and recovery of the mooring in the Drake Passage. This is contribution no. 394 from the Sonderforschungsbereich 95, University of Kiel.

Received 5 April; accepted 7 June 1982.

- Schrader, H.-J. *Science* **174**, 55-57 (1971).
- Hinga, K. R., Sieburth, J. & Heath, G. R. *J. mar. Res.* **37**, 557-579 (1979).
- Cobler, R. & Dymond, J. *Science* **209**, 801-802 (1980).
- Suess, E. *Nature* **288**, 260-263 (1980).
- Thunell, R. & Honjo, S. *Mar. Geol.* **40**, 237-253 (1981).
- Honjo, S., Manganini, S. J. & Cole, J. J. *Deep-Sea Res.* **29**, 609-625 (1982).
- Deuser, W. G., Ross, E. H. & Anderson, R. F. *Deep-Sea Res.* **28**, 495-505 (1981).
- Knauer, G. A., Martin, J. H. & Bruland, K. W. *Deep-Sea Res.* **26**, 97-108 (1979).
- Starecinic, N. *J. mar. Res.* (in the press).
- Wakeham, S. G. *et al. Nature* **286**, 798-800 (1980).
- Deuser, W. G. & Ross, E. H. *Nature* **283**, 364-365 (1980).
- Knauer, G. A. & Martin, J. H. *Limnol. Oceanogr.* **26**, 181-186 (1981).
- Urrere, M. A. & Knauer, G. A. *J. Plankton Res.* **3**, 3-19 (1981).
- Zeitzschel, B., Uhlmann, L. & Diekmann, P. *Mar. Biol.* **45**, 285-288 (1978).
- Tilzer, M. & von Bodungen, B. *Berichte aus dem Institut für Meereskunde* No. 80 (1981).
- Whitworth, T., Nowlin, W. D. Jr & Worley, S. J. *J. Phys. Oceanogr.* (in the press).
- Suess, E. & Ungrer, C. A. *Oceanol. Acta* **4**, 151-160 (1980).
- Berger, W. H. & Killingley, J. S. *Science* **197**, 563-566 (1977).
- Berger, W. H., Killingley, J. S. & Vincent, E. *Oceanol. Acta* **1**, 203-216 (1978).
- Epstein, S., Buchsbaum, H. A., Lowenstamm, H. A. & Urey, H. C. *Bull. geol. Soc. Am.* **64**, 1315-1325 (1953).
- Craig, H. & Gordon, L. I. *Stable Isotopes in Oceanographic Studies* (ed. Tongiorgi, E.) 9-130 (Consiglio Nazionale delle Ricerche, Rome, 1965).
- Müller, P. J. *Geochim. cosmochim. Acta* **41**, 765-776 (1977).
- Handa, N. & Yanagi, K. *Mar. Biol.* **4**, 197-207 (1969).
- Black, C. A. *Methods of Soil Analysis* Pt II, 1038-1040 (American Society of Agronomy, Madison, 1965).

Extreme longevity of lotus seeds from Pulantien

David A. Priestley*† & Maarten A. Posthumus‡

* Department of Plant Physiology, Agricultural University, Arboretumlaan 4, 6703 BD Wageningen, The Netherlands

† Boyce Thompson Institute, Tower Road, Cornell University, Ithaca, New York 14853, USA

‡ Laboratory of Organic Chemistry, Agricultural University, De Dreijen 5, 6703 BC Wageningen, The Netherlands

In 1923 Ohga¹ first brought to general scientific attention the existence of a cache of viable seeds of East Indian lotus (*Nelumbo nucifera* Gaertn.), contained within an ancient lakebed deposit at Pulantien in southern Manchuria. From the geological and historical evidence available, Ohga suggested a possible age in excess of 400 years^{1,2}. Libby³ radiocarbon dated some of Ohga's material at 1040 ± 210 years BP. This evidence of great antiquity for viable seeds has been controversial⁴⁻⁶. The main hindrance to the resolution of the problem has been the paucity of available Pulantien seeds following the dissipation of Ohga's original collection⁷. Recently we have received four Pulantien seeds, three of which were viable, albeit lacking in vigour. The lipids of the unimbibed seeds were examined and found to be still highly polyunsaturated, suggesting that they had undergone little atmospheric autoxidation. Radiocarbon dating of one of the viable seeds suggested a probable age of about 466 years at the time of germination. This is the oldest viable seed for which an age has been directly determined.

The four seeds supplied to us (designated P1 to P4) were collected in 1952 at Pulantien (Hsinchin, Liaoning Province, People's Republic of China)⁸. Following collection they were stored under herbarium conditions. For comparative purposes, we have also analysed freshly grown lotus seeds (designated W1 to W3) obtained from a colony in Washington D.C. that had been derived from two Pulantien lotus seeds germinated in 1951⁵. Each of the Pulantien seeds had the morphological peculiarities described by Ohga⁹ as distinguishing them from recently grown lotus seeds—a waxless, dark-brown pericarp, slightly pitted and lacking a styler protuberance at the apical end. It should be noted that lotus 'seeds' are technically dry fruits, since the seeds in the strict sense are surrounded by an extremely hard pericarp. This massively thickened structure is highly impervious to water and gases unless it is radically scarified^{1,10}. In our investigations, the extreme basal tip of the pericarp was sawn off to expose a small cavity where the cotyledons had shrunk away from the inner pericarp wall. Approximately 10 mg of dry cotyledonous material was excised for lipid analysis and the remainder of the seed was subjected to a germination test. In all the seeds examined the cotyledons were white, except P3 which had dull yellow cotyledons. Germination was initiated by imbibing in tap water at 20 °C for 15 min. The seeds were then placed semi-submerged on cotton wool and left in darkened conditions at approximately 25 °C. All the seeds (except P3) showed some signs of germination within 72–96 hours. After 96 hours the 'fresh' Washington material had plumules extending 39–44 mm beyond the tip of the seed. The corresponding extensions for the Pulantien seeds were less than 20 mm for P1 and less than 10 mm for P2 and P4. During the same period P3 took up water, though the cotyledons did not maintain turgidity and its plumule showed no evidence of growth. After 96 hours, seeds P3 and P4 were heated in an oven at 110 °C for 2 days to kill and dry the material for radiocarbon dating. The seeds were dated at 340 ± 80 BP (GrN-11012) and 430 ± 100 BP (GrN-11013) respectively, which correspond historically (after calibration against dendro-chronologically dated tree-rings¹¹) to AD 1440–1640 and AD 1410–1620. Seeds P1 and P2 were allowed to continue germinating and, after transplanting, they developed into

apparently normal plantlets possessing three or four leaves, at which stage observations were terminated.

One common hypothesis of seed ageing identifies the oxidation of polyunsaturated fatty acid as a possible primary lesion during dry long-term storage¹². Lipids from the lotus seeds were analysed, therefore, as important indicators of the extent of autoxidation in the tissue. The Washington seeds (W1, W2 and W3 of Table 1), which had all presumably matured under rather similar conditions, had very similar fatty acid contents. The degree of polyunsaturation was high, dienoic (C18:2) and trienoic (C18:3) fatty acids accounting for almost half of the total. The total lipid content of the seed was approximately 1% of dry weight. The Pulantien seeds had a more variable fatty acid content than their Washington counterparts, although their similarity to the latter is obvious (Table 1). The much lower linoleic acid (C18:2) content of P3 and P4 compared to P1 and P2 is unlikely to be due to oxidation during isolation. Linolenic acid (C18:3), which is much more sensitive to oxidative degradation, was still quite evident. Most likely, P3 and P4 matured in warm summers, whereas P1 and P2 developed under colder conditions favouring a higher degree of polyunsaturation. Such effects are known to occur in other species^{13,14}.

The oldest documented age for seed viability is for a lotus seed collected in China in 1705 which germinated in 1942, a lifespan of 237 years^{6,15}. Our own data for the Pulantien material suggests that an even greater longevity for this species is possible and that lotus seeds can maintain their high degree of polyunsaturation despite very long periods of quiescence. The situation in lotus is in contrast to that in maize (*Zea mays* L.), for example, which loses the bulk of its polyunsaturated fatty acid within about 100 years¹⁶. The two factors most in favour of prolonged longevity in the Pulantien seeds are probably the extremely thick pericarp and the generally reducing environment present in the peat within which they were deposited. It should be noted that enzymatic repair mechanisms are unlikely to operate in such seeds, which have a moisture content of approximately 11%¹. The lack of extensive lipid autoxidation in the Pulantien material probably accounts for the relatively well-preserved membrane profiles seen in ultra-structural studies^{17,18}. On the basis of such limited data it is difficult to speculate on the causes of viability loss in seed P3. Some lipid autoxidation may have occurred, but not enough to eliminate linolenic acid (C18:3) or to reduce the total level of polyunsaturation below that of the (viable) seed P4. The yellowing of the cotyledons of P3 is probably significant. In a previous study we noted similarly enhanced yellowing in corn seed that had lost viability during thirty to forty years of long-term storage, although this could not always be convincingly linked to fatty acid autoxidation¹⁹.

As noted by Godwin²⁰ complete credibility for claims of

Table 1 Fatty acid contents of Pulantien and Washington lotus seeds

	% Of total fatty acid						
	16:0	18:0	18:1	18:2	18:3	20:0	22:0
P1	25.0	1.3	11.7	49.2	4.5	2.5	5.7
P2	23.7	1.4	12.9	48.0	4.2	3.4	6.4
P3	34.7	2.4	11.8	39.5	2.6	2.5	6.5
P4	31.5	2.8	17.3	36.3	2.6	2.7	6.7
W1	21.6	2.1	19.6	42.4	5.1	3.5	5.6
W2	21.0	1.4	20.5	41.7	4.7	4.0	6.7
W3	22.7	1.3	20.4	43.0	4.6	3.5	4.4

Lipids were extracted from approximately 10 mg of cotyledonary material using 10 ml of chloroform:methanol (2:1, v/v). The solvent extract was filtered and washed once against 2 ml of 0.9% (w/v) NaCl. Fatty acid methyl esters were prepared according to ref. 24. Analyses were made using a Pye Series 204 gas chromatograph coupled to a VG Micromass 70-70 mass spectrometer system. The separations were on a 1.5 m × 2 mm column packed with 3% Silar 10-C on Chromosorb W-HP (100/120 mesh) using a temperature programme of 140–220 °C at 4 °C min⁻¹ or 120–220 °C at 6 °C min⁻¹. (16:0, palmitic; 18:0, stearic; 18:1, oleic; 18:2, linoleic; 18:3, linolenic; 20:0, arachidic; 22:0, behenic). Seeds P1–P4, Pulantien; W1–W3, Washington.

† Address for correspondence.

extreme seed longevity can come only from dating of the seed itself, preferably of a completely inert part such as the coat⁵. The mere dating of an object or structure associated with the seed is insufficient. With the possible exception of a reputedly 620-year-old *Canna compacta*²¹ seed, all other circumstantial claims for long lifespan are seriously suspect⁶. It is worth noting that not only are the lotus seeds among the oldest viable seeds recorded, but also their tissues must contain some of the longest-lived post-mitotic plant cells known. Although some trees such as *Sequoia sempervirens* and *Pinus longaeva* may attain ages of several millennia, the living cells in such structures are unlikely to exceed in lifespan a few decades at most^{22,23}. The cryptobiotic nature of well-protected dehydrated seeds such as the lotus assures their constituent tissues of a potentially much greater longevity.

We thank Dr Pei-sung Tang (Institute of Botany of the Chinese Academy of Sciences) and Dr Liu Shen-cheng (Beijing Botanical Garden) for providing seeds and Dr Horace V. Wester (National Parks Service, Washington DC) for donating freshly grown lotus seeds. Radiocarbon dating was carried out by Dr W. G. Mook (Laboratory of General Physics, State University of Groningen). We thank Professors Johan Bruinsma and Carl Leopold for encouragement. D.A.P. acknowledges the provision of a fellowship by the Netherlands Organization for the Advancement of Pure Research.

Received 25 March; accepted 15 June 1982.

- Ohga, I. *Bot. Mag., Tokyo* **37**, 87–95 (1923).
- Ohga, I. *Bot. Mag., Tokyo* **41**, 1–6 (1927).
- Libby, W. F. *Science* **114**, 291–296 (1951).
- Godwin, H. & Willis, E. H. *New Phytol.* **63**, 410–412 (1964).
- Wester, H. V. *Hortscience* **8**, 371–377 (1973).
- Justice, O. L. & Bass, L. N. *Principles and Practices of Seed Storage* (USDA Agriculture Handbook No. 506, Washington, D.C., 1978).
- Chaney, R. W. *Gdn. J. N.Y. bot. Gdn* **1**, 137–139 (1951).
- Krishtofovich, A. N. *Priroda, Mosk.* Part 1, 76–78 (1953).
- Ohga, I. *Jap. J. Bot.* **3**, 1–20 (1926).
- Ohga, I. *Am. J. Bot.* **13**, 754–759 (1926).
- Stuiver, M. *Radiocarbon* **24**, 1–26 (1982).
- Villiers, T. A. in *Seed Ecology* (ed. Heydecker, W.) 265–288 (Pennsylvania State University, University Park, 1973).
- Howell, R. W. & Collins, F. I. *Agron. J.* **49**, 593–597 (1957).
- Harris, H. C., McWilliam, J. R. & Mason, W. K. *Aust. J. agric. Res.* **29**, 1202–1212 (1978).
- Anonymous *Nature* **149**, 658 (1942).
- Priestley, D. A., Galinat, W. C. & Leopold, A. C. *Nature* **292**, 146–148 (1981).
- Hallam, N. D. & Osborne, D. J. in *Eighth International Congress on Electron Microscopy, Canberra, 1974*, Vol. 2 (eds Sanders, J. V. & Goodchild, D. J.) 598–599 (Australian Academy of Science, Canberra, 1974).
- Zhukova, G. Ya. & Yakovlev, M. S. *Bot. Zhur.* **61**, 869–872 (1976).
- Priestley, D. A., McBride, M. B., Galinat, W. C. & Leopold, A. C. *Pl. Physiol. (Suppl.)* **67**, 115 (1981).
- Godwin, H. *Nature* **220**, 708–709 (1968).
- Lerman, J. C. & Cigliano, E. M. *Nature* **232**, 568–570 (1971).
- MacDougal, D. T. & Smith, G. M. *Science* **66**, 456–457 (1927).
- Fritts, H. C. *Tree Rings and Climate* (Academic, London, 1976).
- Douce, R. & Lance, C. *Physiol. Vég.* **10**, 181–198 (1972).

Leaf injury on wheat plants exposed in the field in winter to SO₂

C. K. Baker, M. H. Unsworth* & P. Greenwood

Department of Physiology and Environmental Science,
University of Nottingham School of Agriculture, Sutton Bonington,
Loughborough LE12 5RD, UK

The combined stresses of air pollution and low temperatures have been suspected as a cause of difficulty in establishing grasses¹ and trees² in upland areas of northern England. Davison and Bailey³ recently showed in controlled environment studies that exposure to SO₂ predisposed grasses to frost injury, and Keller⁴ has established that exposure to SO₂ of young spruce trees kept in cabinets outdoors over winter enhanced their sensitivity to frost injury the following spring. We report here the first observations of injury of winter wheat exposed in the field to controlled concentrations of SO₂ during the severe winter weather of December 1981 to January 1982. We conclude that the injury was probably caused by cold-stress after SO₂ exposure had rendered the plants cold-sensitive.

A microcomputer-controlled system has been designed to release SO₂ from the upwind edge(s) of a square of pipes so that an area 20 m × 20 m in a normally-managed field of winter wheat can be exposed continuously to elevated concentrations of SO₂ from crop emergence in the autumn to harvest in the summer of the following year. The system, described fully elsewhere⁵, has been operated so that, independent of wind speed and direction, the SO₂ concentration at the centre of the square is maintained at a mean value of 100 p.p.b. (267 µg m⁻³ at 20 °C) above ambient, which is typically 10–20 p.p.b. at this site. Plants closer to the upwind edge are exposed to higher concentrations, typically 400 p.p.b. at 2 m from the gas release lines, whereas plants near the downwind edge experience lower concentrations, typically 20 p.p.b.⁵. Thus exposure in the treated area, other than close to the centre, depends on wind direction. By regarding the exposed area as a grid of 25 squares of 4 × 4 m and by combining measured SO₂ concentrations with a numerical diffusion model, the accumulated pollution dose in each grid square (concentration × time) is calculated by the computer.

The field was sown with winter wheat (*Triticum aestivum* L. cv. Bounty) on 23 September 1981. After emergence, the field exposure system was installed and release of SO₂ began on 27 November when the plants were ~4–5 cm high and had four leaves on the main stem. Exposure to SO₂ was continuous apart from breaks for system development and maintenance. Up to 11 December there were no significant differences between exposed plants and controls which were taken from an area ~60 m distant in which SO₂ concentrations did not differ from background values and where soil properties and microclimate were similar to those in the exposed area. On 11 and 14 December ~20 cm of snow fell, covering the plants completely. Gas release continued during the period of snow cover. On 23 December samples of snow (top few mm) were collected at the exposure and control sites for chemical analysis. The snow thawed on 31 December. During the period 12–30 December the mean daily maximum and minimum air temperatures at a Meteorological Office site ~400 m from the wheat field were 1.2 and –6.4 °C, respectively. After a brief mild period, very cold weather returned on 6 January and ~20 cm of snow fell on 8–9 January 1982, remaining until 16 January. Snow samples were collected on 12 January. In the period 6–15 January, mean daily maximum and minimum air temperatures were –1.7 and –9.2 °C, respectively.

A few days after the rapid thaw on 16 January it was clear that leaves of plants within the exposed zone had substantially more injured area than plants in the rest of the field. These injuries, which did not differ significantly throughout the 20 × 20 m zone, consisted of senescing tissue at the leaf tips; injured leaves appeared waterlogged and translucent. An independent assessment revealed no signs of disease. Table 1 shows the percentage of senescing tissue (lamina length) on leaves 4, 5 and 6 of exposed and control samples harvested on three different dates after the thaw. Data for leaf 3 are omitted because these leaves were senescing rapidly throughout the field at this period. Leaves 4 and 5 were injured significantly more in the exposed area; it seems that the injury on leaf 4 developed after the thaw, but injury on leaf 5 did not increase significantly after it was first recorded. A small amount of injury was observed on leaf 6 which emerged mainly after the thaw. Leaves emerging subsequently showed no injury.

We have examined three hypotheses to account for the injury observed:

(1) Leaves were injured due to 'acid' snow resulting from absorption of SO₂ at the snow surface. Injury would have been most likely when the snow thawed and plants were bathed in an acid solution.

(2) The injury was a response to SO₂ uptake by leaves after the plants had been made susceptible to SO₂ injury by low temperatures.

(3) Plants were injured by low temperatures after exposure to SO₂ had reduced their cold-hardiness.

* Present address: Department of Botany, North Carolina State University, Raleigh, North Carolina 27650, USA.

The 'acid snow' hypothesis is not supported by our observations. Analyses of the two sets of snow samples showed insignificant differences between chemical composition of samples from the SO₂-exposed site and the control site. On 12 January the pH of the top few mm of snow ranged from 5.7 to 7.2 in the exposed area and from 6.4 to 7.6 in the control. Sulphate concentration in the snow was 9.5–14.4 µg ml⁻¹ and 9.3–10.8 µg ml⁻¹ at the exposed and control sites, respectively. Analyses of samples collected on 23 December also showed little difference between pH and sulphate concentration at the two sites. The lack of accumulation of sulphate in the snow at the exposed site is consistent with the low deposition velocity reported for SO₂ to snow⁶.

To investigate whether injury could have been caused to roots by a 'pulse' of acid or concentrated sulphate solution leached by melting snow from previously deposited sulphur on plants or soil, we analysed samples of soil solution collected by suction sampling from porous cups at 15 cm depth on 21 January. There were no significant differences in sulphate concentration in the soil solution between exposed and control areas.

It is more difficult to test retrospectively hypotheses (2) and (3). In support of the hypothesis that SO₂ uptake caused the injury is the observation that leaves harvested from near the centre of the exposed area on 21 January contained substantially more total sulphur than controls (8.4 ± 0.2 mg per g dry wt as opposed to 5.1 ± 0.1 mg per g). Dissolved SO₂ dissociates to phytotoxic sulphite and bisulphite forms⁷ and injury is more likely in cold conditions when oxidation to sulphate and other detoxification mechanisms are slow⁸. However, it is necessary to account for the uniformity of injury over the treated area where the dose experienced was very variable. Table 2 shows the computed SO₂ dose for each grid square during daylight hours in the period from 27 November until the injury was observed, excluding periods of snow cover. This dose, which varies by a factor of four across the exposed area would be expected to correlate well with absorbed SO₂ (ref. 9). Thus, unless there was a threshold sulphur content above which injury occurred independent of further sulphur uptake, or unless only leaf tips were susceptible to injury, it is impossible to explain the uniformity of injury. In addition, the symptoms of injury did not resemble the typical well-defined bleached areas attributed to SO₂ injury¹⁰.

The hypothesis that plants were injured by low temperatures after exposure to SO₂ had reduced their cold-hardiness, seems the most probable explanation for the injury, because at least two known mechanisms of SO₂ action on plants are likely to influence cold-hardiness in ways which could account for the uniformity of injury. Black and Black¹¹ observed that brief exposure of bean leaves to SO₂ concentrations similar to those in our experiment decreased the survival of cells on epidermal strips; survival rates did not depend strongly on concentration above ~80 p.p.b. If SO₂ exposure weakened membranes on wheat leaves, it seems likely that resistance to cold stress would be reduced. Alternatively, it is known that in plants where

Table 2 Cumulative SO₂ dose (p.p.m. h⁻¹) calculated for each 4 × 4 m grid square in the experimental area for the daylight hours of 27 November 1981 to 16 January 1982

51.3	28.5	25.8	26.9	40.4
49.0	20.0	16.0	17.2	31.2
49.5	20.1	16.0	16.9	30.3
51.0	24.1	21.3	21.0	31.7
68.7	54.1	51.1	49.7	52.2

Periods when plants were covered with snow are not included.

photosynthetic rates have been reduced or respiration increased, cold-hardiness may be reduced¹². Black and Unsworth¹³ demonstrated that SO₂ concentration > 50 p.p.b. reduced net photosynthesis in beans, and the reduction was independent of SO₂ concentration at low irradiance. Dark respiration was increased, also independent of SO₂ concentration.

We observed no injury after the first cold period even though the lowest night-time air temperature (−15.3 °C) was comparable with the lowest in the second period (−15.8 °C). Temperatures at the surface would have been even colder than these values. It may be that the intervening warm weather with maximum temperatures of 10–12 °C reduced cold-hardiness. On the other hand, day-time temperatures in January were much lower than in December, therefore a threshold for injury may have been exceeded, perhaps because repair mechanisms or replenishment of assimilates could not proceed at the lower temperature.

As with all post mortems, it is difficult to be conclusive about causes of injury. In particular the lack of replication of our treatment and control sites makes definitive interpretation of the causes of injury difficult. It seems essential, therefore, to test our hypotheses of the mechanisms of plant response by experiments in more controlled environments.

Whereas cereals are likely to recover from injury during winter, the delay in establishing leaf area could reduce yield if growth was limited later in the season, for example, by nutrition, drought or high temperatures. Frost injury enhanced by SO₂ exposure could not only reduce quality of horticultural crops but also, by delaying harvest, could cause large financial losses.

This work was supported by the Department of the Environment. We thank Sue Gregson and Geoff Dollard for chemical analyses and Colin and Valerie Black for helpful discussions.

Received 6 May; accepted 24 June 1982.

1. Bell, J. N. B. & Mudd, C. H. in *Effects of Air Pollutants on Plants* (ed. Mansfield, T. A.) 87–103 (Cambridge University Press, 1976).
2. Lines, R. in *Rep. Forest Research* 1977, 18–19 (HMSO, London, 1977).
3. Davison, A. W. & Bailey, I. F. *Nature* **297**, 400–402 (1982).
4. Keller, T. *Gartenbauwissenschaft* **46**, 170–178 (1981).
5. Greenwood, P. *et al.* *Atmos. Envir.* (in the press).
6. Barrie, L. A. & Walmsley, J. L. *Atmos. Envir.* **12**, 2321–2332 (1978).
7. Puckett, K. J. *et al.* *New Phytol.* **72**, 141–154 (1973).
8. Schiff, J. A. & Hodson, R. C. A. *Rev. Pl. Physiol.* **24**, 381–414 (1973).

Table 1 Injury on wheat plants harvested on three dates at a site exposed to SO₂ and a control site

Date	Site	Leaf no.		
		4	5	6
21/1	Control	12 ± 4	6 ± 3	0
	Exposed	16 ± 4	36 ± 5	0
3/2	Control	8 ± 2	10 ± 5	0
	Exposed	40 ± 5	44 ± 4	2 ± 1
12/2	Control	8 ± 2	6 ± 1	0
	Exposed	40 ± 2	40 ± 2	3 ± 1

Values shown are the mean percentage (± s.d.) leaf length injured for plants harvested on the dates shown from a site exposed to SO₂ and a control site. The first two harvests consisted of 6 plants from a larger sample of 0.25 m of two adjacent rows. The third harvest was 60 plants from each site.

9. Fowler, D. & Cape, J. N. in *Effects of Gaseous Air Pollution in Agriculture and Horticulture* (eds Unsworth, M. H. & Ormrod, D. P.) 3–26 (Butterworths, London, 1982).
10. *Diagnosing Vegetation Injury Caused by Air Pollution* (Air Pollution Training Institute, Research Triangle Park, North Carolina, 1976).
11. Black, C. R. & Black, V. J. *J. exp. Bot.* **30**, 291–298 (1979); *Pl. Cell Envir.* **2**, 329–333 (1979).
12. Alden, J. & Hermann, R. K. *Bot. Rev.* **37**, 37–142 (1971).
13. Black, V. J. & Unsworth, M. H. *J. exp. Bot.* **30**, 473–483 (1979).

Variation in genome size—an ecological interpretation

J. P. Grime & M. A. Mowforth

Unit of Comparative Plant Ecology (NERC), Department of Botany, The University, Sheffield S10 2TN, UK

Previous attempts to explore the significance of variation in genome size have involved comparisons with respect to life history^{1–4}, taxonomic and evolutionary affiliations^{5–9} and geographical distribution^{10,11}. Here we examine variation in the British flora. Large genomes are particularly associated with Mediterranean geophytes and grasses in which growth is confined to the cool conditions of winter and early spring. We suggest that large genomes have evolved under circumstances in which growth is limited by the effect of low temperature on rates of cell division and are part of a mechanism whereby growth at low temperature is achieved by rapid inflation of large cells formed during a preceding warm dry season. Where moisture supply allows growth to occur in the summer, temporal separation of mitosis and cell expansion confers no advantage and the longer mitotic cycle of large cells is likely to restrict rates of development; here the effect of natural selection has been to reduce cell and genome size.

Figure 1 describes the distribution with respect to genome size of various classes of flowering plants represented by 162 native or naturalized species of the British flora. Contrary to earlier observations^{1,11}, no consistent difference is apparent

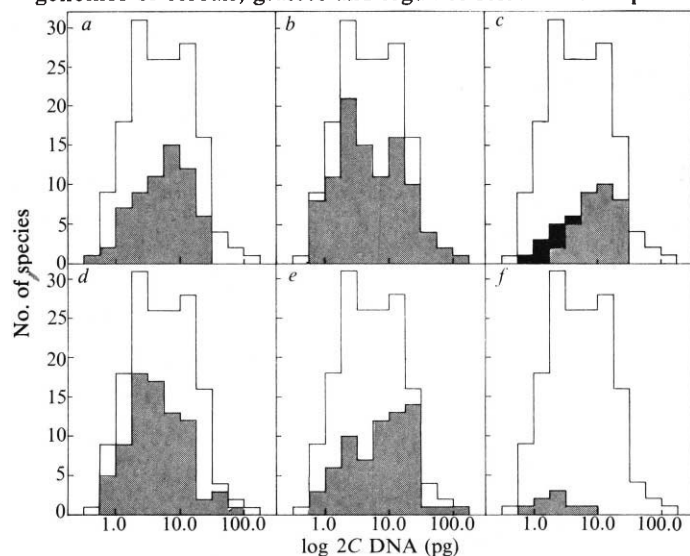
between annuals and perennials. This is due to inclusion in the analysis of common perennial herbs, trees and shrubs, many of which (for example, *Epilobium hirsutum*, *Chamaenerion angustifolium*, *Salix caprea*, *Eriophorum vaginatum*) have comparatively small genomes. Grasses include many species with large (2C DNA > 10 pg) genomes whereas low values are more typical of woody plants. Monocotyledons (mean 2C DNA = 17.28 pg) have significantly larger genomes ($P < 0.001$) than dicotyledons (mean 2C DNA = 6.86 pg). If comparison is restricted to herbaceous plants (Table 1) the difference in mean genome size between monocotyledons and dicotyledons is still observed but no consistent differences are associated with life history. When maps¹² are used to classify the species it is evident that plants found in southern Britain have markedly larger genomes. The mean genome size for the 69 southern species (mean 2C DNA = 13.31 pg) is significantly greater ($P < 0.05$) than that for the 80 species of widespread distribution (mean 2C DNA = 7.96 pg). Although data are available for only eight species, it is interesting that the mean value for plants of northern distribution is low (mean 2C DNA = 2.55 pg).

The species combining southern distribution and large genome size in Fig. 1 include a high proportion of geophytes and grasses with geographical ranges centred on the Mediterranean region. In most of these species, growth occurs during cool winter conditions and survival during the dry summer involves dormancy of both seeds and established plants. Similar phenologies are also common among the species of large genome size and widespread distribution in the British Isles. Here, however, the early development of species such as *Endymion non-scriptus*, *Ranunculus ficaria* and *Caltha palustris* occurs in relatively moist habitats and appears to provide a mechanism of evading competition from productive summer-green trees and herbs¹³.

The correlation between large genome size and the capacity for growth at low temperature is consistent with the enlarged genomes of cereals, grasses and legumes selected for superior

Fig. 1 Genome size in 162 flowering plants occurring in the British Isles.

a, Annuals; b, perennials (including biennials); c, grasses (stippled) and trees and shrubs (shaded); d, widespread distribution; e, southern distribution; f, northern distribution. The distribution for all 162 species is also reproduced in each subsection. In the following key to species and 2C DNA values (pg), original data (obtained by Feulgen microdensitometry on 10 nuclei from each of three seedlings) are distinguished by * from that collated by Bennett and Smith². *Achillea millefolium** (14.4), *Agrostis tenuis** (6.6), *Alliaria petiolata* (3.9), *Allium ursinum** (63.0), *Anthoxanthum odoratum** (11.8), *Anthyllus vulneraria** (1.0), *Aphanes arvensis* (1.1), *Arabidopsis thaliana* (0.5), *Arrhenatherum elatius** (16.0), *Artemisia absinthium* (7.3), *Arum maculatum* (21.8), *Avena fatua* (28.3), *Avena ludoviciana* (27.5), *Beta maritima* (2.6), *Betula pubescens** (1.4), *Brachypodium pinnatum** (2.3), *Brassica napus* (3.2), *Brassica nigra* (1.6), *Brassica oleracea* (1.8), *Briza media* (16.9), *Bromus commutatus* (21.7), *Bromus erectus** (22.6), *Bromus interruptus* (16.8), *Bromus lepidus* (15.7), *Bromus madritensis* (9.7), *Bromus pseudohominii* (19.7), *Bromus racemosus* (27.0), *Bromus rigidus* (17.2), *Bromus sterilis* (6.7), *Bryonia dioica* (3.3), *Caltha palustris* (33.0), *Campanula rotundifolia* (4.9), *Capsella bursa-pastoris* (1.4), *Cardamine flexuosa** (1.6), *Carex caryophylla** (1.5), *Chamaenerion angustifolium** (0.7), *Convallaria majalis* (49.3), *Crepis biennis* (16.9), *Crepis capillaris* (4.7), *Dactylis glomerata* (9.8), *Daucus carota* (2.5), *Digitalis purpurea** (2.3), *Drosera intermedia* (2.0), *Drosera rotundifolia* (1.8), *Endymion non-scriptus* (42.4), *Epilobium hirsutum** (0.6), *Eranthis hyemalis* (18.6), *Eriophorum angustifolium** (1.2), *E. vaginatum** (1.0), *Festuca ovina** (9.5), *Fritillaria meleagris* (141.4), *Fumaria muralis* (1.1), *Helianthemum chamaecistus** (4.5), *Helleborus niger* (22.8), *Holcus lanatus** (3.4), *Hordeum jubatum* (21.7), *Hordeum marinum* (11.0), *Hordeum murinum* (11.1), *Hordeum secalinum* (22.4), *Ilex aquifolium** (2.2), *Impatiens glandulifera** (2.2), *Lactuca serriola* (3.7), *Lamium album* (2.2), *Lamium purpureum* (2.2), *Lathyrus aphaca* (13.8), *Lathyrus maritimus* (15.6), *Lathyrus montanus* (16.5), *Lathyrus nissolia* (13.0), *Lathyrus sylvestris* (23.0), *Lathyrus tuberosus* (18.6), *Leontodon autumnalis* (2.7), *Leontodon hispidus** (5.2), *Lilium pyrenaicum* (65.5), *Lolium italicum* (9.8), *Lolium perenne* (9.9), *Lolium temulentum* (13.6), *Lotus corniculatus** (2.2), *Lotus tenuis* (1.0), *Lotus uliginosus** (1.1), *Lupinus arboreus* (1.6), *Luzula multiflora* (1.9), *Luzula pilosa* (0.6), *Luzula sylvatica* (1.6), *Matricaria matricarioides* (4.6), *Matricaria recutita* (7.8), *Medicago lupulina** (1.6), *Melilotus altissima** (2.5), *Mercurialis perennis* (4.7), *Mibora minima* (5.5), *Molineria caerulea** (4.6), *Montia perfoliata* (4.4), *Myosurus minimus* (2.2), *Origanum vulgare** (1.3), *Ornithogalum longibracteatum* (15.2), *Oxalis acetosella* (6.4), *Papaver dubium* (9.0), *Papaver rhoeas* (5.2), *Papaver somniferum* (7.6), *Petasites hybridus** (1.8), *Phalaris arundinacea* (11.4), *Phalaris canariensis* (7.7), *Phleum bertolonii* (3.4), *Picris echioides* (2.4), *Picris hieracioides* (3.1), *Pimpinella saxifraga** (9.5), *Plantago lanceolata** (2.3), *Poa annua** (3.8), *Poa infirma* (2.4), *Poa trivialis* (5.6), *Ranunculus acris* (10.7), *Ranunculus arvensis* (12.2), *Ranunculus auricomus* (18.0), *Ranunculus baudotii* (8.4), *Ranunculus bulbosus* (11.2), *Ranunculus ficaria* (18.6), *Ranunculus flammula* (12.7), *Ranunculus hederaceus* (4.2), *Ranunculus lingua* (50.2), *Ranunculus ophioglossifolius* (11.3), *Ranunculus paludosus* (16.5), *Ranunculus parviflorus* (12.5), *Ranunculus repens* (22.4), *Ranunculus scardus* (6.4), *Ranunculus scleratus* (8.0), *Ranunculus trichophyllus* (9.7), *Ranunculus tripartitus* (9.0), *Rhinanthus minor* (7.9), *Rumex acetosa** (3.1), *Rumex obtusifolius** (2.7), *Salix caprea** (0.8), *Sambucus nigra** (2.9), *Sarothamnus scoparius** (1.7), *Senecio aquaticus* (3.6), *Senecio squalidus* (1.8), *Senecio vulgaris* (3.0), *Scirpus sylvaticus** (1.3), *Sieglingia decumbens** (5.9), *Silene dioica** (5.0), *Sonchus asper* (3.7), *Stellaria media* (2.1), *Taraxacum officinale* (2.6), *Thymus drucei** (2.6), *Trifolium arvense** (1.4), *Trifolium campestre** (0.9), *Trifolium medium** (6.7), *Trifolium pratense** (1.3), *Trifolium repens** (3.0), *Tripleurospermum maritimum* (5.3), *Tussilago farfara** (4.6), *Urtica dioica** (2.9), *Urtica urens* (0.6), *Veronica persica* (1.5), *Vicia bithynica* (9.1), *Vicia cracca** (11.4), *Vicia hirsuta** (7.2), *Vicia lathyroides* (5.2), *Vicia lutea* (14.8), *Vicia orobus* (10.7), *Vicia sativa* ssp. *angustifolia** (4.0), *Vicia sepium* (9.4), *Vicia sylvatica* (17.2), *Vicia tetrasperma* (7.4).



yield at high latitudes or elevations¹⁰ and with the low values reported for cacti² and tropical plants^{2,9,11}. It is at first surprising that the genomes measured in native plants of northern distribution in Britain (for example, *Eriophorum* spp., *Drosera rotundifolia*) are small. However, at high latitudes winter and spring are severe and unpredictable and the vegetation tends to be dominated by species in which both shoot growth and germination are delayed until the onset of warmer summer conditions^{14,15}.

The massive genomes reported for *Fritillaria* spp. (2C DNA = 96.5–254.8 pg) and *Tulipa* spp. (2C DNA = 39.9–90.6 pg)² are associated with the capacity for early and rapid expansion of relatively short-lived shoots and it is noticeable that their development precedes that of the spring grasses of more moderate genome size with which they are often associated in meadows and gardens respectively. Figure 2 provides a more objective test of the relationship between genome size and phenology. Measurements of the seasonal change in the shoot biomass of species established in natural herbaceous vegetation of the Sheffield region^{13,16,17} have been used to relate the timing of shoot growth in the spring to the genome sizes of individual species. Figure 2 includes all species for which quantitative measurements of both shoot phenology and genome size are available. The results indicate that from early spring to mid-summer changes in the identity of the most actively growing species are associated with a progressive reduction in genome size.

We suggest that the selective force determining this relationship between genome size and season of growth arises from a

Table 1 Comparison of mean genome size (pg) in various classes of herbaceous plant of the British flora

	Annual		Perennial	Total
Monocotyledons	14.07 ***	NS	18.10 **	17.29 ***
Dicotyledons	5.76	NS	8.17	7.17

Probability of significant difference *** <0.001, ** <0.01. NS, not significant.

differential effect of low temperature on cell division and cell expansion. As temperature rises during the spring the advantage of growth dominated by cell enlargement is likely to give way to that of growth involving high rates of cell division. Evidence supporting this hypothesis is available from studies¹⁸ in which mitosis has been shown to be inhibited at temperatures which permit high rates of cell expansion. Variation in genome size is implicated in this proposed mechanism by both our own unpublished observations which confirm that species which habitually grow by cell enlargement at low temperatures have relatively large cells and measurements^{19–23} establishing that potentially large cells have large nuclear volumes and high DNA contents. The mechanism dictating the strong correlation between cell size and nuclear volume remains obscure, although it has been suggested^{3,4} that large cells require an enlarged nuclear envelope to maintain adequate rates of nucleocytoplasmic transport of RNA. It is also possible that cells growing and functioning at low temperatures require more enzymes to sustain adequate metabolic rates. This may necessitate cell enlargement to accommodate extra protein, and higher DNA content could arise from 'dosage repetition'²⁴ of genes coding for proteins and ribosomes²⁵.

As the minimum duration of the mitotic and meiotic cycles is inescapably long in large cells^{19–23} the advantage of growth by cell enlargement would not be expected to pertain at warm temperatures. Thus, small genomes are observed^{2,26} not only in species of continuously warm tropical or arid environments but also in plants in which growth is restricted to the summer period in temperate, continental or arctic zones.

In the geophytes of very large genome size (such as *Fritillaria*, *Tulipa*, *Endymion*), growth occurs mainly by expansion of cells formed during warm relatively dry conditions of the preceding year²⁷. A similar, if less extreme, temporal separation of cell division and expansion may occur in grasses such as *Briza media* and *Anthoxanthum odoratum*, which in Fig. 2 follow the geophytes in the phenological sequence. Here spring growth involves the expansion of relatively large cells but it is less determinate and includes some cell division. In these species, we suggest that spring development may be sustained by the capacity for growth by cell expansion during intermittent cold periods and in the cooler parts of the daily temperature cycle. We suspect also that an integral part of this mechanism may be the accumulation of unexpanded cells during periods in which cell expansion but not cell division is restricted by moisture stress.

Patterns of variation in genome size may be useful in predicting plant response to climate, in crop breeding and in the analysis of niche-differentiation in plant communities. The data compiled by Britten and Davison²⁸ suggest that temperature may have played a major part in the determination of genome size in animals also. As might be expected, small genomes are characteristic of warm-blooded animals and it is particularly interesting that reptiles which flourish in dry hot conditions have uniformly small genomes whereas amphibians include species with exceptionally large genomes and cells. The respiratory system of amphibians depends on the maintenance of a moist permeable skin, so the ecology and behaviour of most species involve avoidance of insulation and maintenance of low body temperature. We suspect that particularly in amphibians exploiting temporary water bodies in a Mediterranean climate, there may be some temporal separation of cell division and expansion; in these conditions natural selection acting on rates of develop-

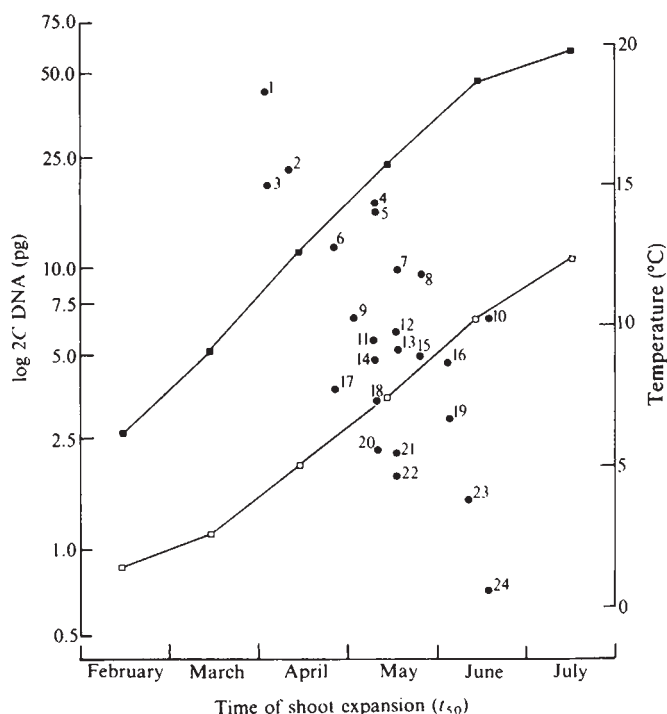


Fig. 2 Relationship between genome size and time of shoot expansion in plant species commonly found in the Sheffield region (correlation coefficient significant at <0.001). The time of shoot expansion is calculated as the date to the nearest quarter-month at which 50% of the maximum shoot biomass was attained^{13,16,17}. Temperature in Sheffield is expressed as the long-term averages for each month of daily minima (□) and maxima (■) in air temperature 1.5 m above the ground surface. 1, *E. non-scriptus*; 2, *R. repens*; 3, *R. ficaria*; 4, *B. media*; 5, *A. elatius*; 6, *A. odoratum*; 7, *L. perenne*; 8, *F. ovina*; 9, *T. medium*; 10, *A. tenuis*; 11, *S. decumbens*; 12, *P. trivialis*; 13, *L. hispidus*; 14, *C. rotundifolia*; 15, *M. perennis*; 16, *M. caerulea*; 17, *P. annua*; 18, *H. lanatus*; 19, *U. dioica*; 20, *B. pinnatum*; 21, *L. corniculatus*; 22, *P. hybridus*; 23, *C. caryophylla*; 24, *C. angustifolium*.

ment may have followed a course similar to that proposed for vernal plants.

We thank Drs M. D. Bennett, D. Goldstein, T. T. Elkington,

Received 13 April; accepted 2 July 1982.

- Bennett, M. D. *Proc. R. Soc. B* **181**, 109–135 (1972).
- Bennett, M. D. & Smith, J. P. *Phil. Trans. R. Soc. B* **274**, 227–274 (1976).
- Cavalier-Smith, T. *J. Cell Sci.* **34**, 247–278 (1978).
- Cavalier-Smith, T. *Bio Systems* **12**, 43–59 (1980).
- Rothfels, K., Sexsmith, E., Heimburger, M. & Krause, M. O. *Chromosoma* **20**, 54–74 (1966).
- Chooi, W. Y. *Genetics* **68**, 195–211 (1971).
- Smith, J. P. & Bennett, M. D. *Heredity* **35**, 231–239 (1975).
- Price, H. J. *Bot. Rev.* **42**, 27–52 (1976).
- Jones, R. N. & Brown, L. M. *Heredity* **36**, 91–104 (1976).
- Bennett, M. D. *Envir. exp. Bot.* **16**, 93–108 (1976).
- Levin, D. A. & Funderburg, S. W. *Am. Nat.* **114**, 784–795 (1979).
- Perring, R. H. & Walters, S. M. *Atlas of the British Flora* (Nelson, Edinburgh, 1962).
- Al-Mufti, M. M., Sydes, C. L., Furness, S. B., Grime, J. P. & Band, S. B. *J. Ecol.* **65**, 759–791 (1977).

B. C. Jarvis, D. L. Lewis, R. Law and D. Watts, and Messrs J. P. Smith, S. Band and C. Mathieson for advice and encouragement. This research was supported by the NERC.

- Billings, W. D. & Mooney, H. A. *Biol. Rev.* **43**, 481–529 (1968).
- Wein, R. W. & Maclean, D. A. *Can. J. Bot.* **51**, 2509–2513 (1973).
- Al-Mashhadani, Y. D. thesis, Univ. Sheffield (1979).
- Furness, S. B. thesis, Univ. Sheffield (1980).
- Haber, A. H. & Luippold, H. J. *Pl. Physiol.* **35**, 168–173 (1960).
- Bradley, M. V. *Am. J. Bot.* **41**, 398–402 (1954).
- Van't Hof, J. & Sparrow, A. *Proc. natn. Acad. Sci. U.S.A.* **49**, 897–902 (1963).
- Commoner, B. *Nature* **202**, 960–968 (1964).
- Darlington, C. D. *Cytology* (Churchill, London, 1965).
- Bennett, M. D. *Proc. R. Soc. B* **178**, 277–299 (1971).
- Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1053–1063 (1977).
- Long, E. O. & Dawid, I. B. A. *Rev. Biochem.* **49**, 727–764 (1980).
- Avdulov, J. P. *Trudy prikl. Bot. Genet. Selek.* **44**, Suppl. 4, 1–428 (1931).
- Hartsema, A. M. *Handb. PflPhysiol.* **16**, 123–167 (1961).
- Britten, R. J. & Davidson, E. H. *Q. Rev. Biol.* **46**, 111–133 (1971).

Competition of similar and non-similar genotypes

J. M. Pérez-Tomé

Colegio Universitario Arcos de Jalón, Arcos de Jalón s/n, Madrid-32, Spain

M. A. Toro

Departamento de Genética Cuantitativa, Instituto Nacional de Investigaciones Agrarias, Carretera de la Coruña km 7, Madrid-35, Spain

It has been well established that the fitness of a genotype is often a function of the presence and relative frequencies of other genotypes coexisting with it^{1–3}. Facilitation or interference between individuals of like genotypes is suggested by some experiments^{4–9}, and it is commonly accepted that competition is an important element in agricultural practice^{10,11}. However, most such studies, at least those that use *Drosophila* as the experimental organism, involve strains which carry different mutations or chromosomal rearrangements. Here we demonstrate that competition of similar as opposed to non-similar genotypes can be detected in a random mating population not subject to genetic manipulation.

In the experiments reported here we used a population of *Drosophila melanogaster* founded from a cross of 70 isofemale lines derived from flies collected by E. Torroja in Carbonera (Almería) in September 1981. Since then the population has been maintained on Lewis food medium. The flies were kept in an incubator room at 25 ± 1.5 °C under continuous illumination. All handling was at room temperature using ether anaesthesia.

Virgin males and females were collected within 6 h of emergence and placed in plastic or glass vials containing *Drosophila* medium, in groups of 20–30 of the same sex. The matings were set up when the flies were 3 days old by placing each one of 80 males with 10 females in a vial for 96 h, after which the male was removed. Two series of vials were then established. In the first, called the homogeneous series, the 10 females that had mated with the same male were placed in a 9 × 3 cm vial with 2 ml of medium. In the second, called the heterogeneous series, 10 females chosen randomly, but excluding any females

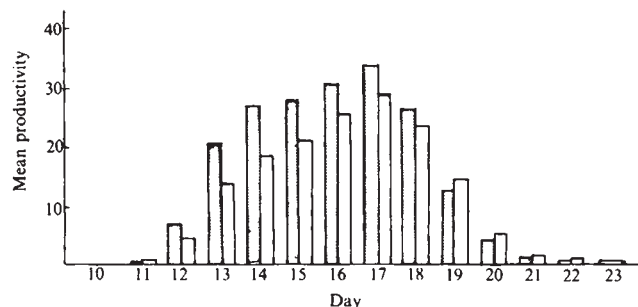


Fig. 1 Mean number of flies emerging daily for the homogeneous (□) and heterogeneous (▨) series.

mated with the same male, were placed in the same vial. All the females were transferred to fresh vials every 24 h for a total of four egg-laying periods. The progeny emerging in each vial were removed and counted daily until the cultures were exhausted.

The only difference between the homogeneous and the heterogeneous series is that in the first the full-sib families that compete are related as half-sibs whereas in the second these full-sib families are randomly assorted. However, the genotypes sampled in the two series are the same, thus explanations such as heterozygote advantage can be disregarded when interpreting the experiment.

Table 1 shows the mean productivity (number of flies emerging per vial) for the different egg-laying periods and the series. Figure 1 shows the daily mean productivity for the two series. This productivity is clearly greater in the heterogeneous than in the homogeneous series; and the difference between the total productivity of the flies in the two series is highly significant ($P < 0.005$).

The results presented here demonstrate clearly that in highly competitive conditions larval viabilities of *D. melanogaster* depend on the similarity of the genotypes coexisting. The productivity can be decreased by 17% merely by modifying the way in which the full-sib families compete.

There are two basically different mechanisms by which some larvae may affect the survival of others: (1) by diffusible metabolic products that affect the survival of other larvae; and (2) by differential utilization of niches by the different genotypes.

Table 1 Mean number of *Drosophila* emerging per vial for the different egg-laying periods and experimental series

Series	Period				Total
	1	2	3	4	
Heterogeneous	58.46 ± 3.16	53.62 ± 2.41	42.62 ± 1.93	37.49 ± 2.04	192.20 ± 5.3
Homogeneous	51.88 ± 3.41	39.88 ± 3.33	38.51 ± 3.00	32.86 ± 2.35	159.69 ± 9.90
<i>t</i>	1.43	3.38	1.16	1.49	2.94
<i>P</i>	<0.20	<0.001	<0.30	<0.20	<0.005

Approximately 35 vials were examined per period for each series; 10 flies per vial. Values shown are mean ± s.d.

It is not clear whether interference or facilitation is involved in these mechanisms.

These results have a bearing on the importance of frequency-dependent selection as an explanation for the maintenance of polymorphism, and are also relevant to the competition between genotypes important in agricultural practice. If this phenomenon is a general one it would also provide a mechanism for the maintenance of recombination through sib-competition¹². The simplicity of the experimental design presented here should allow us to explore more easily the nature of competition between genotypes.

We thank B. Charlesworth, John Maynard Smith, C. Lopez-Fanjul, L. Silvela, J. Malpica and E. Torroja for helpful advice and criticism.

Received 7 June; accepted 5 July 1982.

1. Ayala, F. J. & Campbell, C. A. *A. Rev. Ecol. Syst.* **5**, 115–138 (1974).
2. Lewontin, R. *The Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).
3. Clark, B. *Proc. R. Soc. B205*, 453–474 (1979).
4. Dobzhansky, T. & Spassky, B. *Genetics* **29**, 270–290 (1944).
5. Lewontin, R. *Evolution* **9**, 27–41 (1952).
6. Mather, K. & Cooke, P. *Heredity* **17**, 381–407 (1962).
7. Lewontin, R. & Matsu, Y. *Proc. natn. Acad. Sci. U.S.A.* **49**, 270–278 (1963).
8. Seaton, A. P. C. & Antonovics, J. *Heredity* **22**, 19–33 (1967).
9. Caligari, P. D. S. *Heredity* **45**, 219–231 (1980).
10. Donald, C. M. *Adv. Agron.* **15**, 1–118 (1963).
11. Allard, R. W. & Adams, J. *Am. Nat.* **103**, 621–645 (1969).
12. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).

Single cortical neurones have axon collaterals to ipsilateral and contralateral cortex in fetal and adult primates

Michael L. Schwartz & Patricia S. Goldman-Rakic

Section of Neuroanatomy, Yale University School of Medicine, New Haven, Connecticut 06510, USA

In the adult brain, cortical neurones that project to other cortical areas have traditionally been classified as either 'associational' (having connections within the ipsilateral hemisphere) or 'callosal' (projecting to areas in the contralateral hemisphere). Using double-labelling with fluorescent dyes^{1,2}, we have now identified a novel category of mature cortical neurone that has both an 'associational' and a 'callosal' axon. Such neurones can direct their callosal axons either to a cytoarchitectonic area homotopic to the one in which their cell bodies reside or to an entirely different heterotopic region in the contralateral hemisphere. These cortical neurones with divergent axon collaterals in the adult neocortex differ from recently described neurones that have two axons only transiently during development^{3–5}.

Our evidence is based on data from two fetal (embryonic ages E132 and E137 respectively), two infant (4 and 12 days old respectively) and two adult rhesus monkeys. Each monkey received injections of one dye [Fast blue (FB), Nuclear yellow (NY) or propidium iodide (PI)] in the dorsal bank of the principal sulcus in the left hemisphere and injections of a second dye in the posterior bank of the intraparietal sulcus of the right hemisphere. In addition, one of the monkeys was injected with a third dye in the posterior cingulate cortex of the right hemisphere. Twenty to 72 h after the last injection, the animals were killed and their brains were immediately sectioned and prepared for microscopic examination. The distribution and morphology of labelled callosal neurones were determined in the principal sulcus, the intraparietal sulcus and the cingulate cortex (see outlined areas, Fig. 1). All three regions of association cortex receive projections from each other^{6–11} as well as from the corresponding homotopic area in the opposite hemisphere^{6,7,12}.

Over 95% of fluorescent neurones in the areas examined were single-labelled with FB, NY or PI in all animals (Fig. 2a, b, e). Many of these were callosal neurones, retrogradely filled by dye injections in the corresponding cortical region of the opposite hemisphere. In agreement with previous studies^{10,13–15}, callosal neurones were concentrated in layer III, although some were also found in layers V and VI. All other single-labelled cells, also concentrated in layer III, were association neurones. For example, neurones in the principal sulcus were retrogradely filled by dye injections into the intraparietal cortex in the same hemisphere. Association and callosal neurones were generally intermixed rather than segregated into distinct laminar or radial groupings.

The remaining fluorescent neurones ($\leq 5\%$) were double-labelled and their percentage was no greater in fetuses or infants than in adults. This small population could be divided into two classes: (1) homotopic callosal neurones, which project one axon across the callosum to the same cytoarchitectonic cortex in the contralateral hemisphere (for example, left principal sulcus to the right principal sulcus) and a second axon to a distant cortical target in the same hemisphere (such as left principal sulcus to left intraparietal sulcus) (Fig. 2c, d); and (2) heterotopic callosal neurones which, like their homotopic counterparts, also have axon collaterals in both hemispheres. However, in this case, the target in the contralateral hemisphere is an area different from the homotopic site of origin. For example, double-labelled neurones were found in both anterior and posterior cingulate cortex in each hemisphere, indicating that a cingulate neurone can project simultaneously to the frontal cortex of one hemisphere and the parietal cortex of the other. In one animal injected with PI in the posterior cingulate cortex, several retrogradely labelled PI cells were found in the ipsilateral anterior cingulate cortex; some of them also contained FB which had been injected in the prefrontal cortex of the opposite hemisphere (Fig. 2e, f). Thus, these anterior cingulate neurones projected to both the posterior cingulate cortex of one hemisphere and the prefrontal cortex of the other.

Recent studies of cortical sensory areas in kittens³ and neonatal rats^{4,5} have suggested that, at early stages of postnatal

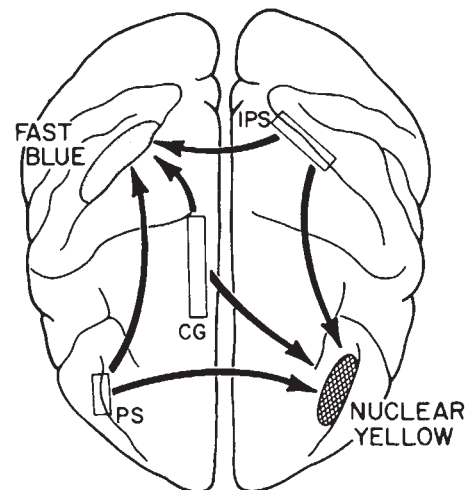
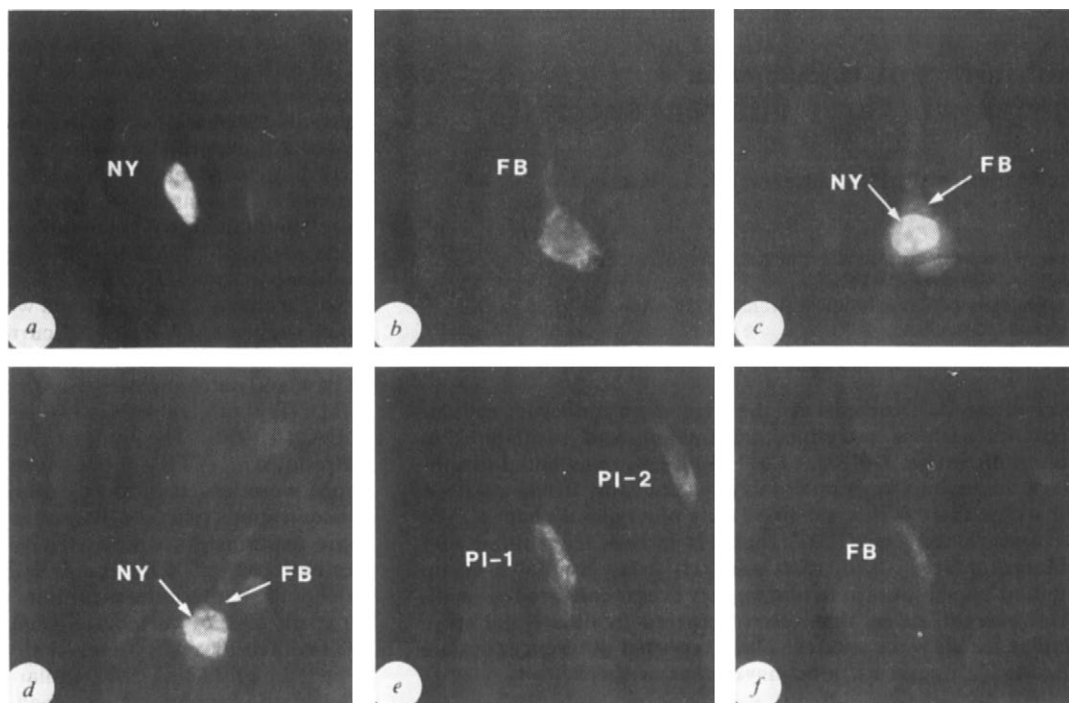


Fig. 1 The basic experimental procedure. Three monkeys received multiple 0.5- μ l injections of 5% FB in the posterior bank of the intraparietal sulcus (IPS) of the right hemisphere and multiple 0.5- μ l injections of 1% NY in the dorsal bank of the principal sulcus (PS) in the left hemisphere. One animal was injected with the same dyes but in reversed locations. The remaining two animals received injections of 3% PI and 5% FB, respectively, in the principal sulcus of the left hemisphere and 1% NY and 3% PI in the intraparietal sulcus, respectively, of the right hemisphere. Shaded areas represent injection sites; areas outlined by open rectangles indicate principle regions examined for cell labelling. Arrows indicate direction of projection of labelled neurones.

Fig. 2 Single- and double-labelled neurones photographed on a Leitz Orthoplan fluorescent microscope equipped with filter cubes D and M. With filter cube D both NY and FB can be simultaneously visualized in the nucleus and cytoplasm, respectively, of retrogradely labelled neurones¹. Filter cube M allows the visualization of PI in the cytoplasm of labelled cells². *a*, Layer III cell in the dorsal bank of the principal sulcus single-labelled with NY injected in contralateral principal sulcus (excitation wavelength 390 nm). Note that only the nucleus is labelled with this dye. *b*, Neurone single-labelled with FB following injection of FB in ipsilateral intraparietal sulcus. With this dye only the cytoplasm is labelled. *c*, *d*,



Double-labelled neurones in layer III from principal sulcus, where the nucleus shows intense yellow labelling with NY while the cytoplasm is labelled blue by FB. These neurones have both a callosal (contralateral) axon collateral to the homotopic principal sulcus and a divergent associational (ipsilateral) axon collateral to the intraparietal sulcus. Note that no significant NY labelling of glia was observed in the areas examined, indicating that, at the survival times used, the NY dye did not leech from labelled cells for re-uptake by nearby neurones and glial cells. *e*, Neurones situated in the anterior cingulate cortex labelled with PI following an injection in the posterior cingulate cortex of the same hemisphere (excitation wavelength 550 nm). *f*, The same field as in *e*, photographed at a different excitation wavelength (390 nm), showing that the neurone labelled PI-1 in *e* is labelled with both PI and FB. Thus, this double-labelled neurone projects to the contralateral principal sulcus and the ipsilateral posterior cingulate cortex.

development, a cortical neurone with intrahemispheric connections issues a callosal collateral which is later lost by selective elimination. In the present study, the density and distribution of double-labelled cortical neurones was the same in fetal, newborn and adult monkeys. It should be emphasized that our injections in monkey fetuses were made 1 month before birth at about the time that cortical axons penetrate the cortical plate in this species¹⁶. Thus, the difference between our findings and those previously reported in other animals cannot be easily attributed to differences in the stages of development examined. A more plausible explanation for the discrepancy between the present results and those of others³⁻⁵ is that there may be a genuine difference in the development of cortical connectivity in acallosal cortical areas such as primary visual cortex, on the one hand, and richly callosal regions of association and limbic cortex, on the other. In either case, our findings clearly demonstrate that for some cortical neurones, collateralization is not a transient developmental state but rather a permanent feature of their adult morphology.

The functional significance of the mature collateralized cortical neurones is not known. Obviously, the simultaneous influence of a single neurone on cytoarchitecturally distinct areas in the two hemispheres is an efficient means for coordinating the activity and transfer of information between widely divergent territories in the two halves of the brain. Such neurones could play a part in interhemispheric transfer by which information is made available to both hemispheres at once^{17,18}. In the present study less than 5% of all neurones were double-labelled, but the territories injected are only a fraction of the total callosal and intrahemispheric targets of the examined areas. Furthermore, although we did not find any triple-labelled neurones in the monkey with injections of three dyes, it is possible that some neurones have more than the two branches identified here. Indeed, future studies may reveal that the two classes of double-labelled neurones described here are actually

members of a single class of callosal neurone with multiple projections to both homotopic and heterotopic callosal targets plus one or more associational targets. Thus, the number of cortical neurones and the number of axon collaterals connecting regions within the same and opposite hemispheres may be considerably larger than indicated by the present results. Accordingly, these results provide new information for interpreting the functions, development and plasticity of the cerebral cortex and alter traditional views of cortical circuitry based on the existence of separate populations of associational and callosal neurones.

This work was supported by USPHS grants NS16666 and MH00298 and fellowship award MH08308. We thank Professor H. G. J. M. Kuypers and Dr Heinz Loewe for providing NY and FB.

Received 10 May; accepted 6 July 1982.

1. Kuypers, H. G. J. M., Bentivoglio, M., Catsman-Berrevoets, C. E. & Bharos, A. T. *Expl Brain Res.* **40**, 383-392 (1980).
2. Kuypers, H. G. J. M., Bentivoglio, M., Van Der Kooy, D. & Catsman-Berrevoets, C. E. *Neurosci. Lett.* **12**, 1-7 (1979).
3. Innocenti, G. M. *Science* **212**, 824-827 (1981).
4. Killackey, H. P. & Ivy, G. O. *Neurosci. Abstr.* **7**, 178 (1981).
5. O'Leary, D. D. M., Stanfield, B. B. & Cowan, W. M. *Dev Brain Res.* **1**, 607-617 (1981).
6. Goldman, P. S. & Nauta, W. J. H. *Brain Res.* **122**, 393-413 (1977).
7. Goldman-Rakic, P. S. & Schwartz, M. L. *Science* **216**, 755-757 (1982).
8. Chavis, D. A. & Pandya, D. N. *Brain Res.* **117**, 369-386 (1976).
9. Jacobson, S. & Trojanowski, J. Q. *Brain Res.* **132**, 209-233 (1977).
10. Hedreen, J. C. & Yin, T. C. T. *J. comp. Neurol.* **197**, 605-621 (1981).
11. Baleyrier, C. & Mauguier, F. *Brain* **103**, 525-554 (1980).
12. Jacobson, S. & Trojanowski, J. Q. *Brain Res.* **132**, 235-246 (1977).
13. Jones, E. G. & Wise, S. P. *J. comp. Neurol.* **175**, 391-438 (1977).
14. Mesulam, M. M., Van Hoesen, G. W., Pandya, D. N. & Geschwind, N. *Brain Res.* **136**, 393-414 (1977).
15. Tigges, J. et al. *J. comp. Neurol.* **202**, 539-560 (1981).
16. Goldman-Rakic, P. S. *The Organization of Cerebral Cortex* (eds Schmitt, F. O., Worden, F. G., Dennis, S. G. & Adelman, G.) 69-97 (MIT Press, Cambridge, 1981).
17. Sperry, R. W. *Science* **133**, 1749-1759 (1961).
18. Gazzaniga, M. S. *Scient. Am.* **217**, 24-29 (1967).

Similarity of unitary Ca^{2+} currents in three different species

A. M. Brown*, H. Camerer, D. L. Kunze* & H. D. Lux

Max-Planck-Institut für Psychiatrie, Kraepelinstrasse 2, D-8000 München 40, FRG

* University of Texas Medical Branch, Galveston, Texas 77550, USA

Membrane Ca^{2+} currents are the triggers for numerous cellular activities such as secretion, contraction, and oscillations in neural discharge. Different Ca^{2+} channels are assumed to subserve the various functions¹; this does not apply to Na^+ currents for which there is thought to be only one type of channel. We have examined single Ca^{2+} channels in cells from three very different species (bird, snail and rat) using the patch-clamp method², and report here that unitary events consisted of small, brief current pulses that often occurred in bursts and were similar for all three species. Thus, reported differences among various Ca^{2+} currents¹ must have other origins.

In our experiments, invertebrate neurones were represented by identified pacemaker cells³ of *Helix pomatia*, vertebrate neurones by cultured dorsal root ganglion (DRG) cells from chick (*Gallus gallus*)⁴, and vertebrate secretory cells by the PC12 line of cultured rat (*Rattus rattus*) adrenal medullary tumour cells^{5,6}. The cultured neurones have macroscopic Ca^{2+} currents that are similar to those of snail neurones^{3,4,7}. Snail neurones were isolated as described elsewhere³ and DRG cells

were prepared using a method that results in glia-free preparations⁸. At least 24 h elapsed before these cells were studied. PC12 cells grown according to Lucas *et al.*⁹ were studied 4–8 h after sub-culture. During experiments, cultured cells were viewed with phase-contrast optics through an inverted microscope. All experiments were done at 20–22°C. Patch-clamp recordings were made using cell-attached patches and isolated patches in the inside-out configuration². The snail neurones were simultaneously voltage-clamped using a two-microelectrode method^{3,10}. Whole-cell clamping with the patch pipette was done in some DRG and PC12 cells.

Single channel Ca^{2+} currents were separated using solutions which suppressed Na^+ and K^+ currents. Osmotic strengths were 250 and 320 mosM for snail and vertebrate cells, respectively. In bath and patch pipettes, Na^+ and K^+ ions were replaced with (mM): Tris 60–100 or Cs 60–100, and tetraethylammonium (TEA) 10–60. The anions were Cl^- and/or gluconate. Tetrodotoxin (TTX), 2–5 μM and 4-aminopyridine (4-AP), 5 mM were added in most experiments. Ca^{2+} was present in concentrations of 20 or 40 mM in the patch pipette, except in some experiments where 20 mM BaCl_2 replaced CaCl_2 ; pH was buffered to 7.4 by Tris or HEPES (Serva). For inside-out patches the bath solutions contained (mM): Cs gluconate (60–80 for snail; 130 for vertebrate cells), hemi-Mg gluconate 20–40, EGTA (Titriplex VI; Merck) 1.0, and TEA- Cl 10–20; pH was buffered to 7.2 using Tris. Similar 'intracellular' solutions were used in the patch pipettes when whole-cell recordings from vertebrate cells were made. In the case of snail neurones, TEA was electrophoresed³ to produce estimated intracellular concentrations of 20–50 mM.

Voltage and current data were recorded on analog tape. Currents were low pass-filtered at 1 or 2.5 kHz by a 4-pole active filter and were digitized at 0.15–0.25 ms per point using a PDP 11/34 minicomputer. Control and test records were

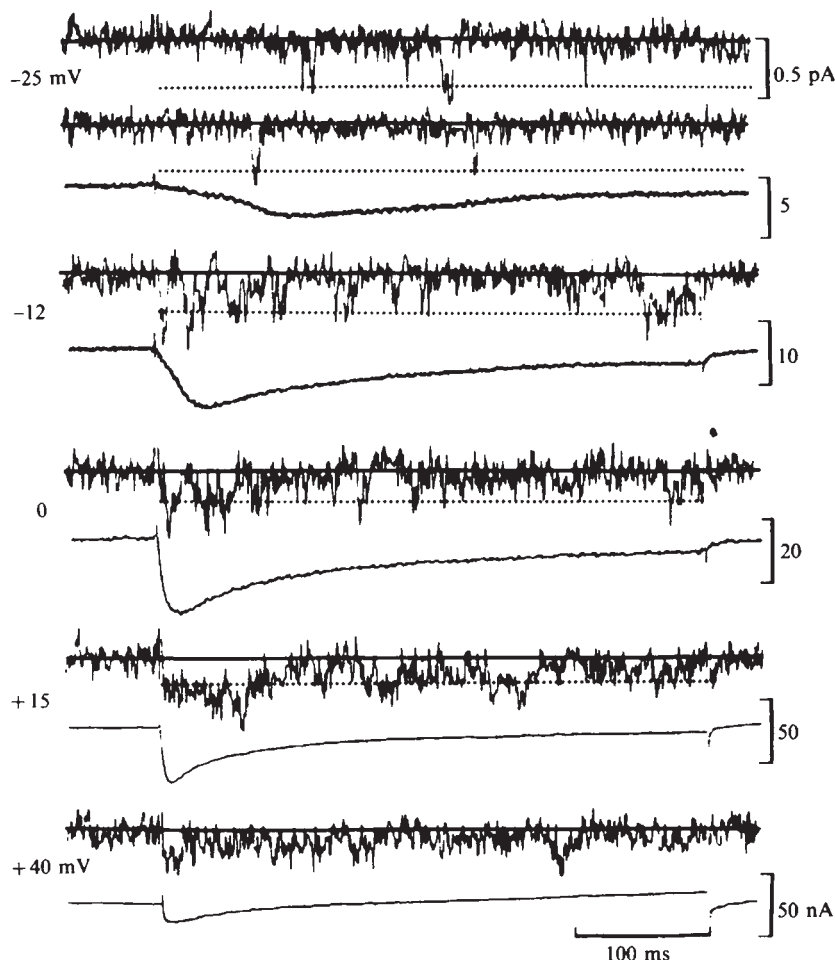
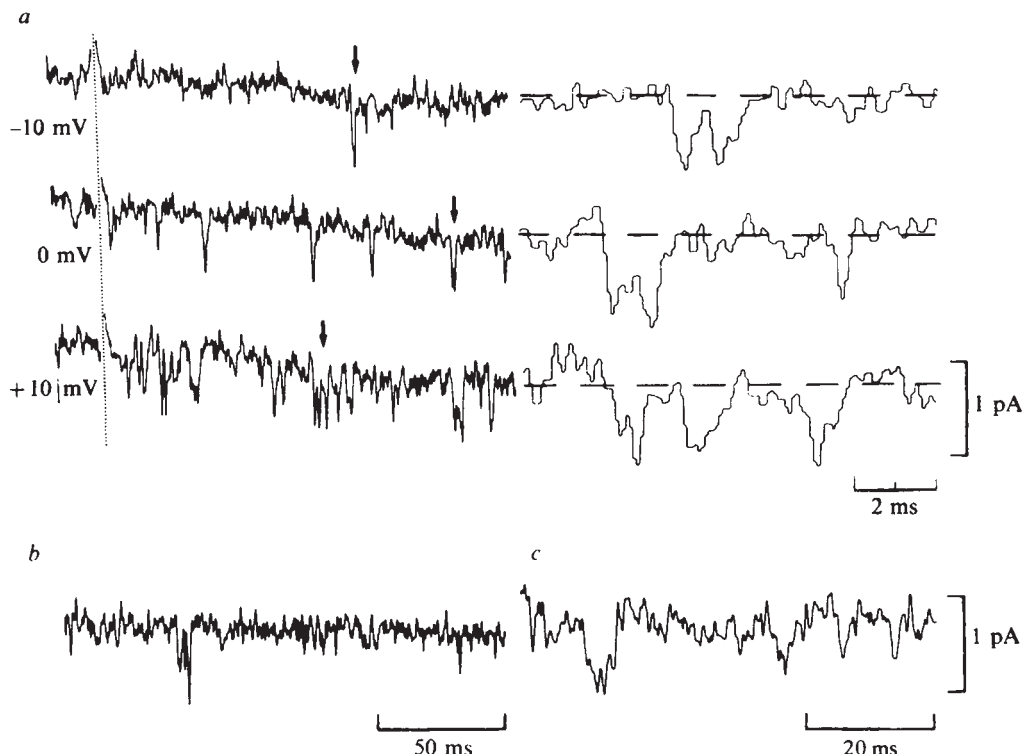


Fig. 1 Unitary membrane Ca^{2+} currents from a snail neurone. The entire neurone was voltage-clamped. Top traces show the patch Ca^{2+} currents and bottom traces the whole-cell membrane Ca^{2+} current during steps to indicated membrane potentials from a holding potential of -50 mV. Linear components of leakage and capacitance have been subtracted. Ca^{2+} concentrations in the patch pipette and bath solution were identical (40 mM). Average unit current amplitude is indicated by dotted lines, and baseline values by the solid lines. As the potential was increased, the frequency of channel openings increased, simultaneous openings from 2 units appeared and unit amplitudes decreased. The data were filtered at 1 kHz.

Fig. 2 *a*, Unitary Ba^{2+} currents from a DRG cell-attached patch at different potentials. The dotted line indicates the time of the voltage step. Components of leakage and capacitance have been subtracted. A resting membrane potential of -50 mV was assumed. The threshold for activation of unitary Ba^{2+} currents was -20 mV. With increasing potential, the frequency of channel openings increased. Openings that occur in a burst were observed. Sections of the current records marked with an arrow in the left-hand panel are shown on an expanded time scale and bandwidth in the right-hand panel; the data were filtered at 1 kHz and 2.5 kHz, respectively. *b*, Unitary Ba^{2+} currents from a PC12 inside-out patch at $+30$ mV. Note a burst of openings with a superimposed single opening. Holding potential, -50 mV. *c*, Unitary Ca^{2+} currents from a DRG inside-out patch at $+20$ mV. Holding potential, -50 mV. Data were filtered at 1 kHz (*b*, *c*). 20 mM Ba^{2+} (*a*, *b*) or Ca^{2+} (*c*) was present in the patch pipette.



0.3–1.0 s long, and sets of at least 12 records were used for unit analysis.

Background noise was 0.1–0.2 pA r.m.s. at d.c. to 1.0 kHz bandwidth. As the signal-to-noise ratio was low, we used two different methods of measuring amplitudes and open and closed times. With the first method an interactive program was used to excise unitary currents that occurred singly or in bursts. Mean values for the background current were computed and used as reference levels. This method required subjective estimates of the unitary currents. A more objective method was to apply threshold levels at $\sim 2/3$ and $1/3$ of the expected current amplitude to mark the beginning and end of a channel current. When necessary, time-varying mean currents were removed before computing the histograms. The different methods were applied to two sets of identical data and were found to give similar results.

In snail neurones, the membrane potential was recorded directly together with macroscopic currents. Unitary patch currents which appeared similar were recorded repeatedly from different parts of the same neurone) this apparently caused no

damage to the cell, although an inward current often appeared transiently on removal of the patch electrode. Membrane potential was not recorded in DRG and PC12 cells. However, the resting potentials^{6,11} were probably similar to the holding potential of -50 mV used for snail neurones.

Unitary currents were observed following $G\Omega$ seals in 13 patches from DRG cells, 16 patches from PC12 cells, and >200 patches from snail neurones. In all cases, these currents were characterized by the appearance of brief current pulses of small amplitude (0.5 ± 0.1 pA) at potentials that were depolarized by 20–30 mV from the resting potential. The openings occurred throughout steps of 1.0 s duration, but appeared to be less frequent at longer times. The openings occurred singly or in bursts (see Figs 1, 2). Long openings without evidence of any closure were also observed. Occasionally, amplitudes of about twice the unit height (Figs 1, 2*b*) were observed; these were interpreted as simultaneous openings of two channels. Openings became more frequent with increased depolarization (Figs 1, 2*a*), and the unitary current amplitudes decreased. In snail neurones the slope conductance was calculated to be 6–7 pS.

Table 1 Ca^{2+} channel characteristics

	Snail neurone	Rat PC12 cell	Chick DRG neurone
Current amplitude $\pm$ s.d. (pA)	-0.47 ± 0.11 (120)	-0.53 ± 0.12 (1115)	-0.53 ± 0.10 (790)
Closed time (ms)			
within bursts	1.3 (15)	0.9 (543)	0.7 (274)
between bursts	46.2 (94)	21.9 (550)	20.7 (484)
Open time (ms)	2.8 (120)	2.7 (1115)	0.5 (790)
Duration of bursts (ms)	3.4 (106)	6.8 (577)	1.5 (516)
No. of openings per burst	1.2 (106)	1.9 (577)	1.5 (516)
k_2 (ms^{-1})	0.64	0.58	0.95
k_3 (ms^{-1})	0.13	0.53	0.48
k_4 (ms^{-1})	0.36	0.37	2.00

Mean values of Ca^{2+} channel characteristics from cell-attached patches of snail neurone (-20 mV), rat PC12 cell (-10 mV) and chick DRG neurone (10 mV). For analysis, we selected recordings which showed a discrete amplitude level on opening, presumably from one channel, and came from 1, 3 and 2 cell-attached patches for snail, rat and chick cells, respectively. The resting potential of DRG and PC12 cells was assumed to be -50 mV; that of snail neurone was -50 mV. Patch pipettes contained 40 mM Ca^{2+} for snail neurone and 20 mM Ba^{2+} for PC12 and DRG cells. The standard deviation of current amplitudes was due almost completely to the background noise current. In the calculation of the duration of bursts and number of openings per burst, single openings have been included. k_2 , k_3 and k_4 are the rate constants estimated according to the three-state model¹⁶. Numbers in parentheses represent the numbers of measurements from which the values shown were calculated. In the case of the closed times, the numbers were estimated from the integrals of the fast and slow components. The amplitude and open time values for the snail neurone patch are similar to values from larger samples, recorded previously¹⁰.

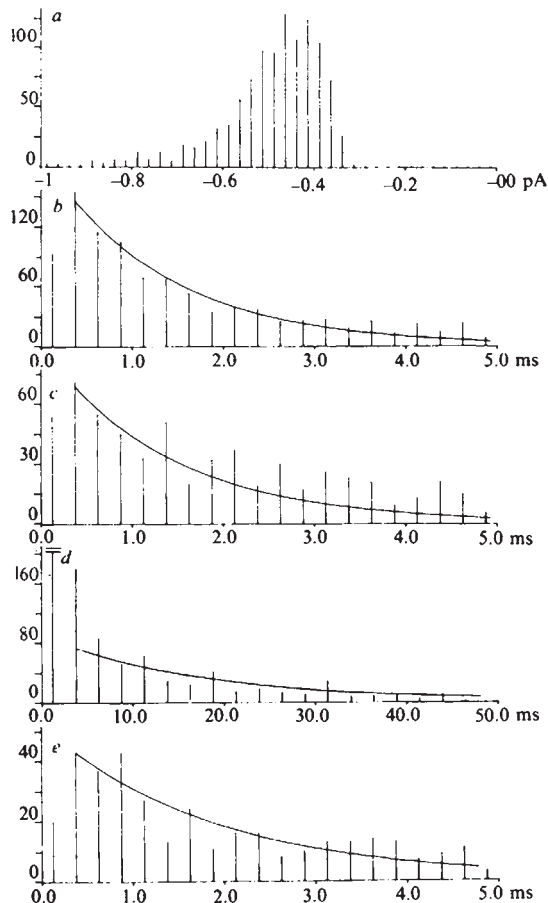


Fig. 3 Histograms of amplitude (a), open time (b), closed time (c, d) and burst duration (e) distributions for unitary Ba^{2+} currents from a PC12 cell-attached patch. Ba^{2+} concentration was 20 mM. The closed time distribution is shown on two different time scales to allow the two exponential components to be distinguished easily. Holding potential was assumed to be -50 mV and the patch potential was stepped to -10 mV. The histograms were computed from 21 records, each of 500 ms duration. Data were filtered at 1 kHz and sampled at 4 kHz. Values were obtained by the threshold method described in the text. The exponential curves in the histograms have time constants of b, 1.3; c, 1.8; d, 17.0; and e, 1.9 ms.

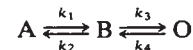
The inward currents became slightly larger when Ba^{2+} replaced Ca^{2+} (Fig. 2a, b). For inside-out patches the pipette interior was made positive by 50–70 mV. In the case of cell-attached patches, channels were activated with depolarizing steps ranging from 30 to 100 mV, and openings became more frequent at larger depolarizations. The currents remained almost unchanged for 10–15 min, which was the usual survival time of the inside-out patches in our solutions.

Table 1 gives the similarities in average amplitudes of the unitary currents for the different cells from the three species examined. Open times for snail neurones and PC12 cells showed similar distributions; the values for DRG cells were shorter, possibly because of differences in resting and activating potentials. The properties of the bursts were similar (see Fig. 3). The open times were described by a single exponential function. The closed times were similar among the different cell groups and could be separated into two exponential functions having short and long time constants. The short time constants are due to the intervals associated with bursts of activity, which were more prominent for Ba^{2+} currents. The longer time constants are due to the intervals between bursts.

The unitary activity could not have been produced by the flow of Ba^{2+} or Ca^{2+} through Na^+ or K^+ channels as the former

were blocked by TTX and the latter by TEA, 4-AP and Cs substitution. TTX-resistant Na^+ channels can be excluded as they do not conduct Ca^{2+} ions¹². The present inward currents were never observed when 10 mM Mg^{2+} or Co^{2+} replaced Ba^{2+} or Ca^{2+} , consistent with observations on macroscopic Ca^{2+} currents. The unitary Ba^{2+} currents were rather larger, a result also predicted from the macroscopic currents⁷. The unitary current amplitudes agree with those reported previously¹⁰ but are larger than values estimated from noise measurements^{13,14}. The present data are insufficient to allow any significant conclusion to be made regarding differences between divalent cations or the effects of changes in divalent concentration.

Our data do, however, indicate that open Ca^{2+} channels have only one conduction level. The bursting activity points to the fact that activation is more complicated than a simple first-order process; see also ref. 10. Further evidence for this is the delayed turn-on of Ca^{2+} currents which can be fitted by an m^2 model¹⁵ and the bi-exponential nature of Ca^{2+} tail currents^{16,17}. The simplest kinetic scheme consistent with our data is



where A and B are closed states and O is open. The rate constants k_2 , k_3 and k_4 can be estimated from the histograms¹⁸ and are given in Table 1. This model was also used by Fenwick *et al.*¹⁷.

The units of Ca^{2+} conduction were similar in the three different species studied here. The unitary currents produced in isotonic Ba^{2+} solutions in bovine chromaffin cells¹⁷ and in neonatal rat heart cells¹⁹ also resemble the units described here. It is possible that, as the cells from two of the species were cultured, the units might not be representative of the original cells. This possibility cannot be excluded, but the fact that the Ca^{2+} currents and action potentials recorded from these cultured cells^{4,11} are very similar to those recorded from a wide range of species and tissues¹ renders such a suggestion unlikely. Hence, we conclude that Ca^{2+} channels everywhere are basically the same. The reported differences among macroscopic Ca^{2+} currents may be due to, among other factors, the variety of receptors associated with Ca^{2+} channels^{4,20} and the second-messenger actions of Ca^{2+} ions²¹.

We thank Drs C. Lucas, Y. Barde and R. Deisz, Ms G. Schuster and Ms S. Messerschmidt, and Mr R. Childers. We especially thank Mr H. Zucker for help with the computer programs and Drs E. M. Fenwick, A. Marty and E. Neher for providing us with a preprint of their paper. A.M.B. was supported by a Josiah Macy Fellowship and NIH grant NS 11453. H.C. was supported by DFG grant HE 1120/2-2.

Received 14 April; accepted 29 June 1982.

- Hagiwara, S. & Byerly, L. A. *Rev. Neurosci.* **4**, 69–125 (1981).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Hofmeier, G. & Lux, H. D. *Pflügers Arch. ges. Physiol.* **391**, 242–251 (1981).
- Dunlap, K. & Fischbach, G. D. *J. Physiol., Lond.* **317**, 519–535 (1981).
- Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2424–2428 (1976).
- O'Laigue, P. H. & Huttner, S. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1701–1705 (1980).
- Brown, A. M., Morimoto, K., Tsuda, Y. & Wilson, D. L. *J. Physiol., Lond.* **320**, 193–218 (1981).
- Barde, Y. A., Edgar, D. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1199–1203 (1980).
- Lucas, C. A., Edgar, D. & Thoenen, H. *Expl Cell Res.* **121**, 79–86 (1979).
- Lux, H. D. & Nagy, K. *Pflügers Arch. ges. Physiol.* **391**, 252–254 (1981).
- Spector, I. in *Excitable Cells in Tissue Culture* (eds Nelson, P. G. & Lieberman, M.) 247–274 (Plenum, New York, 1981).
- Kostyuk, P. G., Veselovsky, N. S. & Tsyndrenko, A. Y. *Neuroscience* **6**, 2423–2430 (1981).
- Akaike, N., Fishman, H. M., Lee, K. S., Moore, L. C. & Brown, A. M. *Nature* **274**, 379–382 (1978).
- Krishtal, O. A., Pidoplichko, V. I. & Shakhvalov, Yu. A. *J. Physiol., Lond.* **310**, 423–434 (1981).
- Kostyuk, P. G., Krishtal, O. A. & Shakhvalov, Yu. A. *J. Physiol., Lond.* **270**, 545–568 (1977).
- Tsuda, Y., Wilson, D. L. & Brown, A. M. *Biophys. J.* **37**, 181a (1982).
- Fenwick, E. M., Marty, A. & Neher, E. *J. Physiol., Lond.* (in the press).
- Colquhoun, D. & Hawkes, A. G. *Proc. R. Soc. B* **211**, 205–235 (1981).
- Reuter, H., Stevens, C. F., Tsien, R. & Yellen, G. *Nature* **297**, 501–503 (1982).
- Yatani, A., Tsuda, Y., Akaike, N. & Brown, A. M. *Nature* **296**, 169–171 (1982).
- Rasmussen, H. & Goodman, D. B. P. *Physiol. Rev.* **57**, 421–509 (1977).

Single-channel currents in isolated patches of plasma membrane from basal surface of pancreatic acini

Y. Maruyama & O. H. Petersen*

The Physiological Laboratory, University of Liverpool, PO Box 147, Brownlow Hill, Liverpool L69 3BX, UK

Precise localization and characterization of conductance pathways in glandular epithelia have so far proved difficult¹. The patch-clamp technique for high resolution current recording²⁻⁴, which has already been applied successfully to a number of electrically excitable cells^{3,4}, can in principle overcome these difficulties. We now report measurements of single-channel currents from isolated patches of plasma membrane (inside-out) from the baso-lateral surface of collagenase-isolated rat and mouse pancreatic acini. We have identified a cation channel having a conductance of ~30 pS and a mean open time in the range 0.3–1 s which is dependent on internal calcium. The single-channel current–voltage relationship is linear and the mean open time independent of the membrane potential. These channels may, at least in part, account for the Ca^{2+} -mediated neural and hormonal control of pancreatic acinar membrane conductance, which is probably responsible for the Ca^{2+} -dependent acinar fluid secretion⁵⁻⁷.

We used collagenase-isolated rat or mouse pancreatic acini prepared as described previously⁸⁻¹⁰. With this procedure, clusters of acinar cells (5–20) retaining intact tight junctions and the normal morphological polarity are obtained⁸⁻¹⁰. Fabrication of patch recording pipettes, isolation of inside-out membrane patches and high-resolution current recording were carried out as described in detail by Neher and collaborators^{3,4,11}. The total Ca^{2+} concentration in the bath and pipette solutions used was measured by atomic absorption spectroscopy.

Figure 1a shows current recordings from a single patch of plasma membrane with Na-sulphate solution on both sides of the membrane. We observed unitary events of constant amplitude at the same numerical membrane potential (–20 and +20 mV), but with a larger amplitude at a higher membrane potential (–40 mV). In the absence of a membrane potential, no unitary steps were detected. Unit current steps of this type were observed in 47 patches from 28 batches of isolated rat pancreatic acini. Similar experiments performed on pancreatic acini from mice revealed the same type of unit current steps (8 membrane patches from 4 batches of acini).

There was a linear relationship between the amplitude of the unitary current steps and the membrane potential, with reversal of current polarity at 0 mV. At zero membrane potential we observed no channel opening or closure using the same NaCl or Na_2SO_4 solutions on both sides. The existence of a Ca^{2+} gradient, typically 1.28 mM in the pipette and 40 μM in the bath, did not change the reversal potential. Figure 1b shows plots of single-channel current–voltage relationships from one series of experiments in which NaCl or Na_2SO_4 solutions (same concentration on both sides of the membrane) were used. As the single-channel conductance was independent of the presence of Cl^- , these channels seem to be cation-selective. Figure 1b also includes data from experiments in which ~50% of the Na_2SO_4 was replaced by K_2SO_4 ; this did not change the conductance. It seems reasonable therefore to assume that the channels do not discriminate very well between Na^+ and K^+ . The mean value of the single-channel conductance in the three experiments carried out in NaCl solutions was 27 pS, whereas in one series with Na_2SO_4 solutions a mean of 33 pS (10 patches) was obtained.

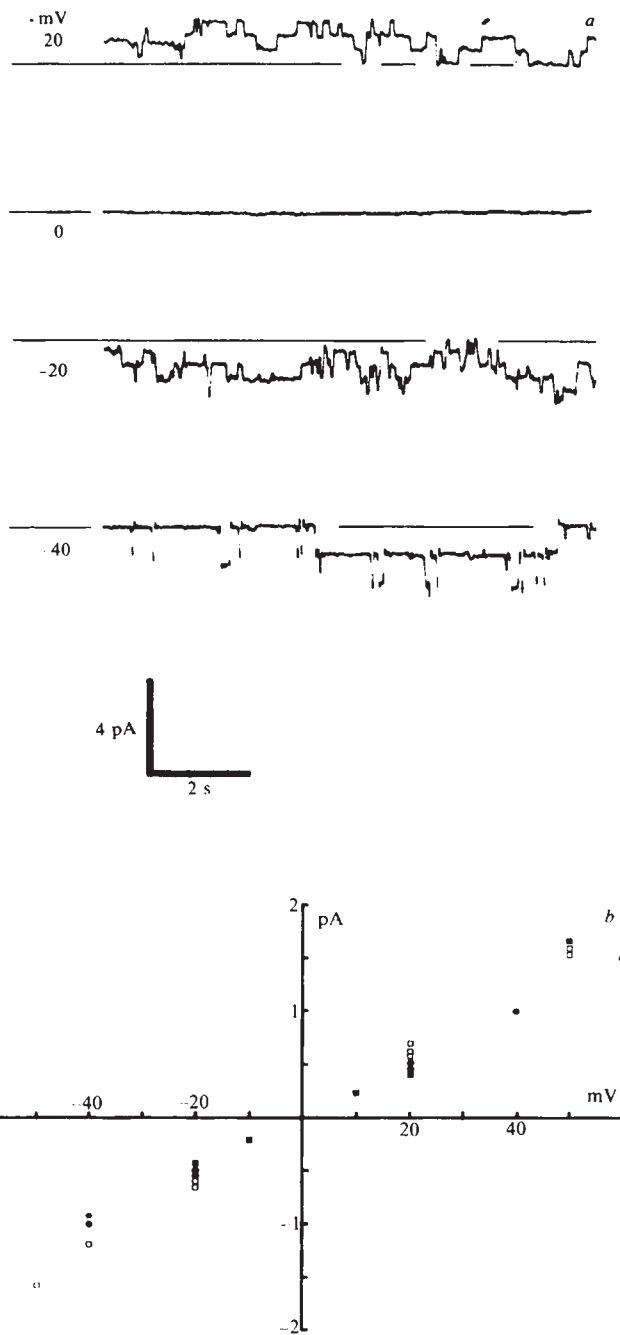


Fig. 1 Single-channel current–voltage relationship. *a*, Recordings of single ion channel currents from isolated membrane patch taken from the baso-lateral surface of collagenase-isolated rat pancreatic acinus. Short segments of the recordings from the same patch at four different membrane potentials are shown. The isolated membrane patch was 'inside-out', that is, the physiological inside of the membrane faced the bath, whereas the outer surface of the membrane faced the interior of the pipette. Outward currents (current from the intracellular to the extracellular side of the membrane) are represented by upward deflections. The membrane potentials indicated are conventional in that, for example, a potential of –40 mV denotes that the physiological inside of the membrane is 40 mV negative with respect to the outside. The patch resistance was 5 G Ω . In this experiment the same solution was present on both sides of the membrane and had the following composition (mM): Na_2SO_4 , 80; KCl, 4.7; CaCl_2 , 1.28; MgCl_2 , 1.15; glucose, 10; HEPES, 10. The pH of the solution (titrated by NaOH) was 7.3 and the temperature 22 °C. 20 Hz filtering (low pass) was applied. *b*, Relationship between the single-channel current and membrane potential. Values shown are from experiments in which (●) a solution with 110 mM NaCl, 10 mM HEPES, 10 mM glucose, 1.15 mM MgCl_2 and 4.7 mM KCl was present on both sides of the membrane (□) a Na_2SO_4 solution (NaCl replaced isosmotically by Na_2SO_4) was present on both sides; and (■), a solution with 40 mM Na_2SO_4 and 45 mM K_2SO_4 was used on both sides. In these three series of experiments the external total Ca^{2+} concentration (pipette fluid) was in the range 1.28–2.56 mM, while the internal total Ca^{2+} concentration (bath fluid) ranged from 40 μM to 1.28 mM.

* To whom correspondence should be addressed.

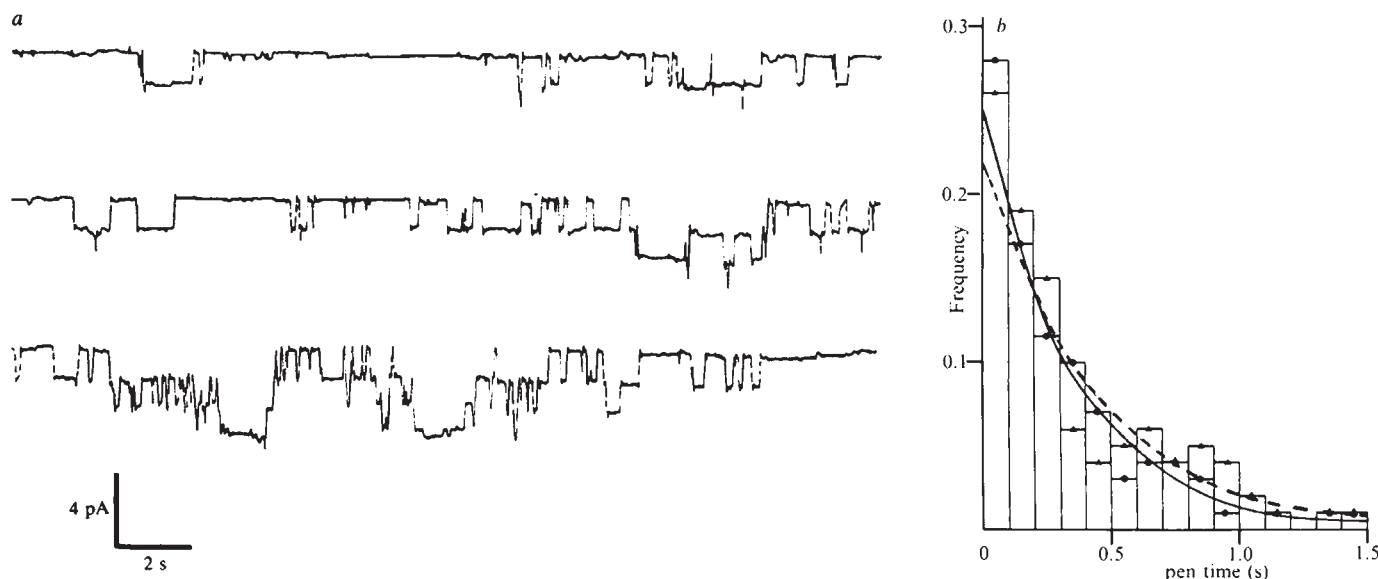


Fig. 2 Temporal characteristics of channel opening and closure. *a*, Continuous trace of single-channel current recording from an inside-out patch exposed to Na-sulphate solution on both sides with 2.56 mM Ca^{2+} in the pipette fluid and 40 μM Ca^{2+} in the bath fluid. The Mg concentration on both sides was 1.15 mM; membrane potential, -50 mV; patch resistance, 15 G Ω . The frequency distributions of both the open and closed times for all the events in the entire recording ($n = 94$ and 77, respectively) were similar to exponential distributions with a single-channel mean open time of 0.93 s and a mean closed time of 0.83 s, if we assume that there are three mutually independent channels and that opening and closure of each of these is a Poisson-type phenomenon. Clamping the membrane potential at 0 mV abolished unitary current steps in this and all other experiments. *b*, Histograms obtained from one membrane patch at membrane potentials of -40 mV ($\bullet$; $n = 108$) and +40 mV ($\blacktriangle$; $n = 110$), showing the relative (normalized) frequency distribution of the open times. The curves ($\bullet$ — $\bullet$) (-40 mV) and ($\blacktriangle$ — $\blacktriangle$) (+40 mV) represent the theoretical exponential distributions for mean open times of 0.36 s and 0.43 s, respectively. In this case the patch apparently only contained one channel.

In one patch for which a relatively long continuous recording was obtained (Fig. 2*a*), the mean single-channel open time was calculated to be 0.93 s whereas the closed time was 0.83 s. We also investigated the effect on the mean open time of changing the membrane potential. In one patch the relative frequency density distribution was compared for membrane potentials of -40 and +40 mV. Figure 2*b* shows the two histograms superimposed on each other. For comparison, the theoretical exponential (Poisson) curves are shown. The two distributions are clearly very similar.

In one series of experiments we investigated the internal calcium dependence of the single-channel currents. Varying the total Ca^{2+} concentration in the bath fluid between 10 μM and 1.25 mM did not markedly affect the current recordings. However, complete removal of bath fluid Ca^{2+} (no Ca^{2+} added, 1 mM EGTA) stopped unitary current activity (Fig. 3). The effect of Ca^{2+} removal was reversible in that the unitary current steps reappeared after Ca^{2+} addition, however, the open time was much shorter and the frequency of opening reduced compared with the first control period. In all five experiments (on five different patches) in which Ca^{2+} was removed from the bath fluid during continued recording, the unitary current steps disappeared.

The channels we have found in the acinar membrane patches seem in many respects similar to the Ca^{2+} -dependent inward current channels recently described in cultured cardiac cells¹² and neuroblastoma cells¹³. This type of channel, so far found in muscle, nerve and gland cells, may be an important feature of many different cells, having a number of roles in various tissues.

The existence of Ca^{2+} -dependent cation permeation pathways in the baso-lateral surface membrane of pancreatic acinar cells was postulated previously on the basis of indirect evidence^{5,14,15}. It is probable that this channel, which has now been directly demonstrated, is important in the stimulus-secretion coupling sequence of acinar cells. Internal Ca^{2+} by simultaneously activating exocytosis and Na^+ entry through

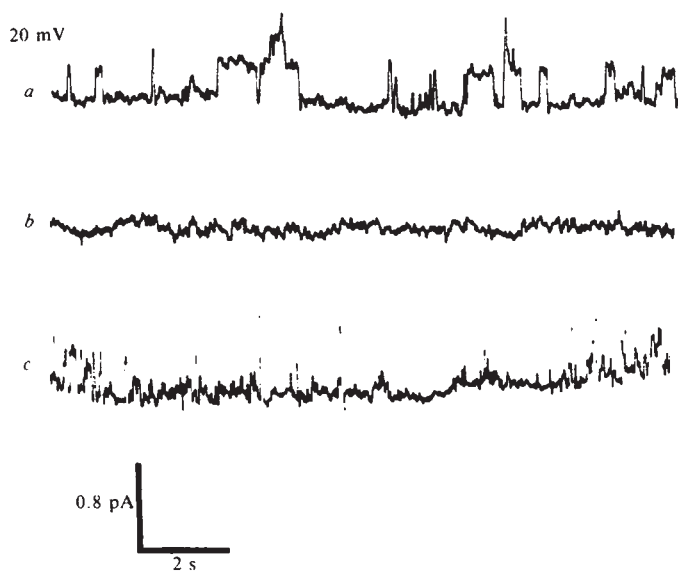


Fig. 3 Ca^{2+} dependency of single-channel currents. *a*–*c*, Consecutive traces from a continuous recording from the same inside-out membrane patch. In this experiment (patch resistance, 3 G Ω) a Na_2SO_4 solution with 1.15 mM MgCl_2 , 4.7 mM KCl, 10 mM glucose and 10 mM HEPES (pH 7.3) was used on both sides. The external (pipette) solution contained 1.28 mM Ca^{2+} . The bathing fluid total Ca^{2+} concentration was 1.28 mM in *a*, while trace *b* was obtained 2 min after changing the bath solution to one without Ca^{2+} but with added EGTA (1 mM). Trace *c* was obtained ~5 min after return of the control solution (1.28 mM Ca^{2+}). The unitary outward current steps seen in *a* disappeared after removal of internal Ca^{2+} , but returned again in *c* after restoration of Ca^{2+} ; however, in this period open times were severely shortened. In other experiments recordings were made of inward current steps at membrane potentials of -20 or -40 mV. Removal of Ca^{2+} from the bath fluid always abolished the unitary current steps.

these Ca^{2+} -dependent channels, will couple the two physiologically important processes in the acinar cells, fluid and enzyme secretion, to hormone-receptor interaction. Several other important elements in the stimulus-secretion coupling sequence still remain to be positively identified, that is, the link between hormone-receptor interaction and activation of the Ca^{2+} -effector^{6,14} and the Cl^- (anion) permeation mechanism and its activation¹⁶. Further work with the patch-clamp technique should help in the characterization of the transepithelial transport mechanism and its control.

We thank Dr E. Neher for practical advice and demonstration of the techniques for high-resolution current recording, and Dr I. Schulz for help in the preparation of isolated pancreatic acini. This work was supported by a project grant from the MRC.

Received 20 May; accepted 11 June 1982.

- Petersen, O. H. *The Electrophysiology of Gland Cells* (Academic, London, 1980).
- Neher, E. & Sakmann, B. *Nature* **260**, 799–802 (1976).
- Neher, E. in *Techniques in Cellular Physiology* (ed. Baker, P. F.) 121, 1–16 (Elsevier, Amsterdam, 1982).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Iwatsuki, N. & Petersen, O. H. *Nature* **268**, 147–149 (1977).
- Petersen, O. H., Nishiyama, A., Laugier, R. & Philpott, H. G. in *Drug Receptors and Their Effectors* (ed. Birdsall, N. J. M.) 63–73 (Macmillan, London, 1981).
- Petersen, O. H. *et al. Phil. Trans. R. Soc. B* **296**, 151–166 (1981).
- Williams, J. A., Korb, M. & Dormer, R. L. *Am. J. Physiol.* **235**, E517–E524 (1978).
- Peikin, S., Rottman, A. J., Batzri, S. & Gardner, J. D. *Am. J. Physiol.* **235**, E743–E749 (1978).
- Wakasugi, H. *et al. J. Membrane Biol.* **65**, 205–220 (1982).
- Neher, E., Sakmann, B. & Steinbach, J. H. *Pflügers Arch. ges. Physiol.* **375**, 219–228 (1978).
- Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752–754 (1981).
- Yellen, G. *Nature* **296**, 357–359 (1982).
- Petersen, O. H. & Iwatsuki, N. *Ann. N.Y. Acad. Sci.* **307**, 599–617 (1978).
- Schulz, I. & Heil, K. *J. Membrane Biol.* **46**, 41–70 (1979).
- Petersen, O. H. & Philpott, H. G. *J. Physiol., Lond* **306**, 481–492 (1980).

Modifier role of internal H^+ in activating the Na^+ – H^+ exchanger in renal microvillus membrane vesicles

Peter S. Aronson, Jeannette Nee & Marjorie A. Suhm

Departments of Medicine and Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

The intracellular pH in animal cells is generally maintained at a higher level than would be expected if H^+ were passively distributed across the plasma membrane¹. In a wide variety of cells including sea urchin eggs², skeletal muscle³, renal and intestinal epithelial cells^{4,6}, and neuroblastoma cells⁷, plasma membrane Na^+ – H^+ exchangers mediate the uphill extrusion of H^+ coupled to, and thus energized by, the downhill entry of Na^+ . Plasma membrane vesicles isolated from the luminal (microvillus, brush border) surface of renal proximal tubular cells possess a Na^+ – H^+ exchanger^{4,5} that seems to be representative of the Na^+ – H^+ exchangers found in other tissues. For example, the renal microvillus membrane Na^+ – H^+ exchanger, like other Na^+ – H^+ exchangers, mediates electroneutral cation exchange^{4,5}, is sensitive to inhibition by the diuretic drug amiloride^{5,8}, and has affinity for Li^+ in addition to Na^+ and H^+ (refs 5, 9). Here we have examined the effect of internal H^+ on the activity of the Na^+ – H^+ exchanger in renal microvillus membrane vesicles. Our results suggest that internal H^+ , independent of its role as a substrate for exchange with external Na^+ , has an important modifier role as an allosteric activator of the Na^+ – H^+ exchanger. Allosteric behaviour with respect to internal H^+ is a property that would enhance the ability of plasma membrane Na^+ – H^+ exchangers to extrude intracellular acid loads and thereby contribute to the regulation of intracellular pH.

Microvillus membrane vesicles were isolated from rabbit renal cortex by a modification¹⁰ of the Mg-aggregation method of Booth and Kenny¹¹. For membranes prepared in this manner, the enrichment in specific activity (final pellet/homogenate) of

luminal membrane markers, such as γ -glutamyl transpeptidase, is >10-fold, whereas those of markers for basolateral membranes and mitochondria, such as $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ and succinic dehydrogenase, respectively, are <1.0 (refs 10, 11). Electron microscopy reveals that this membrane preparation consists predominantly of right-side-out microvilli with intact core structures¹². Studies using freeze-fracture techniques and immunological methods have also demonstrated that >85% of renal microvillus membrane vesicles are oriented right-side-out¹³.

The effect of internal pH on the Na^+ uptake rate is illustrated in Fig. 1. In this experiment, the vesicles were pre-equilibrated (120 min at 20°C) in media of pH 7.47–6.16 and then the 5-s uptake of 1 mM Na^+ was assayed at external pH 7.26. We have previously shown that equilibration of imposed pH gradients occurs within 120 min in these membrane vesicles⁵. The rate of Na^+ uptake was progressively stimulated as internal pH was lowered. This stimulation of Na^+ uptake by internal H^+ was almost completely abolished by 1 mM amiloride, as earlier observed⁸. Amiloride is a specific, competitive inhibitor for the Na^+ – H^+ exchanger in these membrane vesicles⁸. However, the K_i of 3×10^{-5} M for amiloride inhibition of the renal microvillus membrane Na^+ – H^+ exchanger is considerably higher than the apparent K_i (10^{-8} – 10^{-7} M) for inhibition of Na^+ channels in tight epithelia¹⁴. Consequently, drug concentrations of 10^{-3} M must be used to abolish Na^+ – H^+ exchange in these membrane vesicles⁸. The finding that amiloride-sensitive Na^+ uptake was stimulated by internal H^+ certainly represented the behaviour expected for a Na^+ – H^+ exchange process. If the stoichiometry

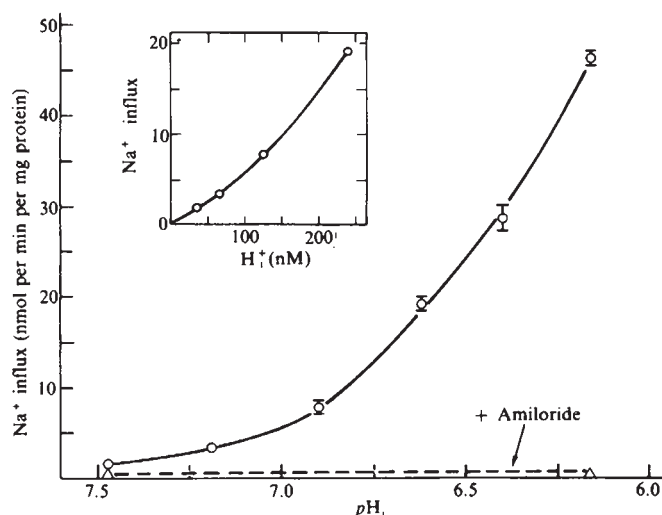


Fig. 1 Effect of internal pH on Na^+ influx rate. Membrane vesicles were isolated and initially suspended in 200 mM mannitol, 41 mM K^+ , 80 mM HEPES pH 7.47, and then pre-equilibrated for 120 min at 20°C in 157 mM mannitol, 74 mM K^+ , 52 mM Cl^- , 42 mM HEPES pH 7.47 or with isosmotic replacement of mannitol by 9, 17, 26, 35 or 52 mM 2-(*N*-morpholino)ethane sulphonate (MES) at pH 7.19, 6.90, 6.62, 6.40 or 6.16, respectively. The membrane vesicles were then incubated for 5 s at 20°C with 1.0 mM ^{22}Na , 45 mM K^+ , 11 mM Cl^- , 178 mM mannitol, 67 mM HEPES, 10 mM MES pH 7.26 in the absence (○) or presence (Δ) of 1.0 mM amiloride. These incubations were performed in a volume of 50 μl, containing 0.1 μCi ^{22}Na (NEN) and 100–200 μg membrane protein. The incubations were terminated by rapid addition of 3.5 ml of an iced (0–4°C) 'stop' solution consisting of 160 mM KCl, 2.5 mM Tris, 4 mM HEPES pH 7.5. The mixture was immediately poured on a 0.65 μm Millipore filter (DAWP) and washed with an additional 10.5 ml of 'stop' solution. Filters were placed in vials containing 3.0 ml Ready Solv HP (Beckman) and counted by scintillation spectroscopy. Values for the nonspecific retention of radioactivity by the filters were subtracted from the values for the incubated samples. Each datum represents the mean ± s.e. for three determinations. The inset shows a plot of Na^+ influx versus internal H^+ concentration for the pH range 7.47–6.62.

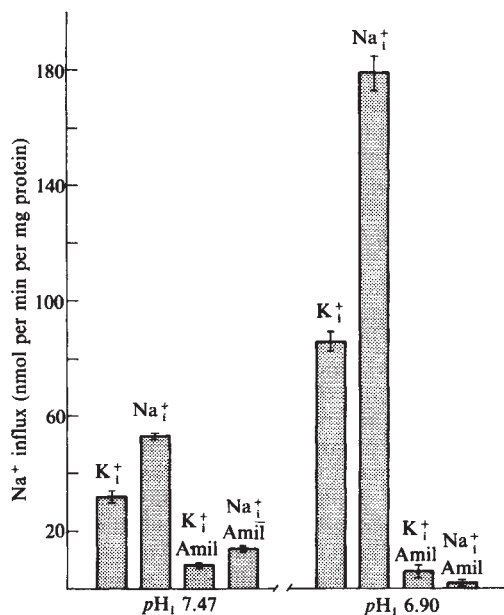


Fig. 2 Effect of internal pH on Na⁺ influx via Na⁺-Na⁺ exchange. Membrane vesicles were pre-equilibrated for 120 min at 20 °C either in 157 mM mannitol, 22 mM K⁺, 42 mM HEPES and 52 mM KCl (K⁺) or 52 mM NaCl (Na⁺) at pH 7.47; or in 140 mM mannitol, 22 mM K⁺, 42 mM HEPES, 17 mM MES and 52 mM KCl (K⁺) or 52 mM NaCl (Na⁺) at pH 6.90. Using the same filtration method as for Fig. 1, the 5-s uptake of ²²Na was then assayed in the presence of 10 mM ²²Na, 43 mM K⁺, 21 mM Cl⁻, 175 mM mannitol, 63 mM HEPES, 3.5 mM MES pH 7.39 with ('Amil') or without 2.0 mM amiloride. Each datum represents the mean ± s.e. for three determinations.

of exchange were 1 H⁺:1 Na⁺, one would expect the Na⁺ uptake rate to increase linearly with increasing internal H⁺ (first order kinetics) and then to plateau (zero order kinetics) as the internal H⁺ transport site became saturated with H⁺. As indicated in the inset to Fig. 1, however, in the pH range 7.47–6.62 the Na⁺ uptake rate has a greater than first order dependence on the internal H⁺ concentration. This finding raised two possibilities. First, the Na⁺-H⁺ exchanger might possess more than one transport site for internal H⁺. Given that Na⁺-H⁺ exchange in these vesicles seems to be electroneutral^{4,5} with a H⁺:Na⁺ coupling ratio¹⁵ of 1.0, this would imply that the stoichiometry of Na⁺-H⁺ exchange, rather than 1:1, is 2:2, 3:3, etc. However, arguing against this possibility, the dependence of the Na⁺ uptake rate on the external Na⁺ concentration conforms to simple Michaelis-Menten kinetics^{8,9}. Second, the Na⁺-H⁺ exchanger might possess only a single transport site for internal H⁺ but, in addition, possess one or more modifier sites at which internal H⁺ could bind and thereby activate the exchanger without being transported. Therefore, additional experiments were performed to attempt to demonstrate that internal H⁺, independent of its role as a transportable substrate for exchange, could actually activate the Na⁺-H⁺ exchanger.

Previous studies have suggested that the renal microvillus membrane Na⁺-H⁺ exchanger can mediate Na⁺-Na⁺ exchange⁹. In the experiment shown in Fig. 2, we examined the effect of internal pH on the rate of ²²Na uptake occurring via exchange for internal Na⁺. Membrane vesicles were pre-equilibrated with 52 mM KCl or 52 mM NaCl at pH 7.47 or 6.90, and then the 5-s uptake of 10 mM ²²Na was assayed at external pH 7.39. Because the renal microvillus membrane Na⁺-H⁺ exchanger has no detectable affinity for K⁺ (refs 5, 9), ²²Na uptake into vesicles containing K⁺ should occur only by Na⁺-H⁺ exchange, whereas ²²Na uptake into vesicles preloaded with unlabelled Na⁺ could occur by both Na⁺-H⁺ exchange and Na⁺-Na⁺ exchange. Thus, at any given internal pH, the difference between the ²²Na influx rates of vesicles preloaded with Na⁺ and those preloaded with K⁺ should represent the

component of ²²Na uptake occurring via Na⁺-Na⁺ exchange. As indicated in Fig. 2, the difference between the ²²Na influx rates of Na⁺-preloaded and K⁺-preloaded vesicles was much greater at internal pH 6.90 than at 7.47. That is, internal H⁺ stimulated the component of ²²Na uptake occurring via Na⁺-Na⁺ exchange. This internal H⁺-stimulated Na⁺-Na⁺ exchange was amiloride sensitive, consistent with the concept that it represented a mode of operation of the Na⁺-H⁺ exchanger. Accordingly, this experiment strongly suggested that internal H⁺, independent of its transport role, could indeed activate the Na⁺-H⁺ exchanger. Figure 2 appears to indicate that the amiloride-insensitive Na⁺ influx was not the same in all conditions. However, the actual values for Na⁺ uptake in the presence of amiloride were less than twice the nonspecific retention of ²²Na by the Millipore filters, probably accounting for the observed variability. No consistent relationship between amiloride-insensitive Na⁺ uptake and either internal H⁺ or internal Na⁺ was clearly evident on replications of this experiment.

The renal microvillus membrane Na⁺-H⁺ exchanger can mediate net Na⁺ transport in either direction^{8,15}. When [Na⁺]_o/[Na⁺]_i exceeds [H⁺]_o/[H⁺]_i, there is net influx of Na⁺ in exchange for internal H⁺; when [Na⁺]_i/[Na⁺]_o exceeds [H⁺]_i/[H⁺]_o, there is net efflux of Na⁺ in exchange for external H⁺ (ref. 15). In the experiment illustrated in Fig. 3, we tested the effect of internal pH on the rate of net Na⁺ efflux occurring via exchange for external H⁺. Membrane vesicles were pre-equilibrated with 52 mM Na⁺ at pH 7.47 or 6.90, diluted 1:100 into a Na⁺-free medium at pH 7.47, and the net efflux of Na⁺ was assayed over the subsequent 30-s period. As indicated in Fig. 3, the net efflux of Na⁺ from vesicles with internal pH 7.47. This internal H⁺-stimulated Na⁺ efflux was amiloride sensitive, consistent with the fact that it represented transport via the Na⁺-H⁺ exchanger. Clearly, the thermodynamic driving force for net Na⁺ efflux via Na⁺-H⁺ exchange ($\Delta\mu_{Na^+} - \Delta\mu_{H^+}$)¹⁶ was less favourable at internal pH 6.90 than at 7.47. Accordingly, the findings in Fig. 3 indicated that internal H⁺ had activated the Na⁺-H⁺ exchanger sufficiently to more than compensate for the inhibitory effect of a diminished driving force for net Na⁺ efflux. This experiment thus provided further support for the concept that internal H⁺ was a potent activator of the renal microvillus membrane Na⁺-H⁺ exchanger.

Because our studies were performed using isolated plasma membrane vesicles rather than intact cells, we can conclude

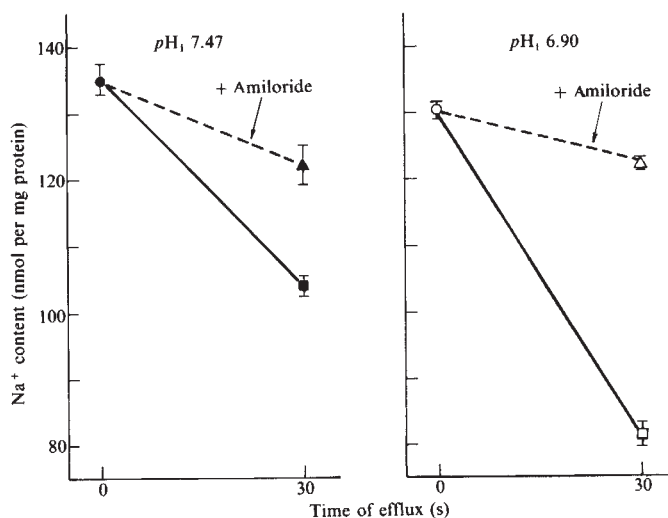


Fig. 3 Effect of internal pH on Na⁺ efflux. Membrane vesicles were pre-equilibrated for 120 min at 20 °C in 52 mM ²²NaCl, 22 mM K⁺, 42 mM HEPES and either 157 mM mannitol at pH 7.47 (●, ▲, ■) or 140 mM mannitol and 17 mM MES at pH 6.90 (○, △, □). Intravesicular ²²Na content was then assayed before (●, ○) and after 1:100 dilution and reincubation of the vesicles for 30 s at 20 °C in 200 mM mannitol, 41 mM K⁺, 80 mM HEPES pH with (▲, △) or without (■, □) 1.0 mM amiloride. Each datum represents the mean ± s.e. for three determinations.

that the observed activation of the Na^+-H^+ exchanger by internal H^+ must have arisen from an interaction at the level of the microvillus membrane. The simplest molecular model to explain our findings is that the Na^+-H^+ exchanger is a transmembrane protein with one or more inwardly facing titratable groups to which internal H^+ can bind and thereby cause a conformational change resulting in an increased transport activity. In experiments not illustrated, we found that the initial rate of Na^+ -dependent glucose uptake in renal microvillus membrane vesicles was not affected by a change in intravesicular pH from 7.47 to 6.90. Thus, whatever the precise molecular mechanism by which internal H^+ activates the Na^+-H^+ exchanger, this activation must not represent a general, nonspecific alteration of microvillus membrane transport function.

In almost all animal cells, a steep inwardly directed Na^+ gradient is maintained across the plasma membrane due to the active extrusion of Na^+ via the $(\text{Na}^+ + \text{K}^+)\text{ATPase}$. Because $[\text{Na}^+]_0/[\text{Na}^+]_i$ exceeds $[\text{H}^+]_0/[\text{H}^+]_i$ in most conditions, the net driving force acting on plasma membrane Na^+-H^+ exchangers generally favours the net exchange of external Na^+ for internal H^+ . Thus, allosteric activation by internal H^+ would enhance

the ability of plasma membrane Na^+-H^+ exchangers to protect intracellular pH against intracellular acid loads. Recent studies suggest that activation of plasma membrane Na^+-H^+ exchangers may also have a role in such other physiological processes as the regulation of cell volume and the control of cell proliferation^{17,18}. It has been proposed that elevations in intracellular Ca^{2+} may induce this activation¹⁸. We have found that preloading renal microvillus membrane vesicles with 10^{-7} – 10^{-3} M Ca^{2+} has no significant effect on Na^+-H^+ exchange activity (Barrett and P.S.A., unpublished observations). Because of mitochondrial $\text{Ca}^{2+}-\text{H}^+$ exchange, perturbations that induce a rise in intracellular Ca^{2+} can cause a fall in intracellular pH¹. Our studies raise the possibility that the apparent activation of plasma membrane Na^+-H^+ exchangers by intracellular Ca^{2+} could actually represent allosteric activation by intracellular H^+ .

We thank Ginger James for secretarial assistance. This work was supported by USPHS Research grant AM-17433 and an Established Investigatorship from the American Heart Association (P.S.A.).

Received 2 April; accepted 7 July 1982.

1. Roos, A. & Boron, W. F. *Physiol. Rev.* **61**, 296–434 (1981).
2. Johnson, J. D., Epel, D. & Paul, M. *Nature* **262**, 661–664 (1976).
3. Aickin, C. C. & Thomas, R. C. *J. Physiol., Lond.* **273**, 295–316 (1977).
4. Murer, H., Hopfer, U. & Kinne, R. *Biochem. J.* **154**, 597–604 (1976).
5. Kinsella, J. L. & Aronson, P. S. *Am. J. Physiol.* **238**, F461–F469 (1980).
6. Rindler, M. J. & Saier, M. H. Jr *J. Biol. Chem.* **256**, 10820–10825 (1981).
7. Moolenaar, W. H., Boonstra, J., van der Saag, P. T. & de Laat, S. W. *J. Biol. Chem.* **256**, 12883–12887 (1981).
8. Kinsella, J. L. & Aronson, P. S. *Am. J. Physiol.* **241**, F374–F379 (1981).

9. Kinsella, J. L. & Aronson, P. S. *Am. J. Physiol.* **241**, C220–C226 (1981).
10. Aronson, P. S. *J. Membrane Biol.* **42**, 81–98 (1978).
11. Booth, A. G. & Kenny, A. J. *Biochem. J.* **142**, 575–581 (1974).
12. Aronson, P. S. & Kinsella, J. L. *Fedn Proc.* **40**, 2213–2217 (1981).
13. Haase, W., Schäfer, A., Murer, H. & Kinne, R. *Biochem. J.* **172**, 57–82 (1978).
14. Cuthbert, A. W. & Shum, W. K. *Proc. R. Soc. B189*, 543–575 (1975).
15. Kinsella, J. L. & Aronson, P. S. *Biochim. biophys. Acta* **689**, 161–164 (1982).
16. Aronson, P. S. *Am. J. Physiol.* **240**, F1–F11 (1981).
17. Kregonow, F. M. *A. Rev. Physiol.* **43**, 493–505 (1981).
18. Benos, D. J. *Am. J. Physiol.* **242**, C131–C145 (1982).

Prostaglandins modulate macrophage Ia expression

David S. Snyder, David I. Beller & Emil R. Unanue

Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA

Prostaglandins are important modulators of inflammation and of humoral and cellular immune responses^{1–8}. In order to evaluate a possible mechanism for the regulation of immune responses we have studied the effects of prostaglandins on the expression of I-region-associated (Ia) antigens by macrophages. The expression of these glycoproteins is essential for macrophages to function as antigen-presenting cells during the induction of immune responses⁹. The synthesis and membrane expression of Ia, however, is not a constitutive property of the phagocyte¹⁰ but is under regulation and a positive regulation of this process is exhibited by activated T cells¹¹. In contrast, a negative regulation is conspicuously found in the neonate where a product from a young replicating macrophage inhibits the expression of Ia by the mature macrophages¹². We show here that prostaglandins of the E series (PGE) are potent inhibitors of the expression of Ia-antigens on macrophages and that thromboxane B₂ (TXB₂) antagonizes the effect of PGE.

The effects of PGE₁ and PGE₂ on the surface Ia expression of cultured macrophages were tested initially. Macrophages were obtained from normal mice or from mice previously injected with protease-peptone or infected with *Listeria monocytogenes*. Resident macrophages from normal mice contained 5 to 10% Ia-positive cells; those from peptone-induced exudates had 5 to 15% positive, while those from *Listeria*-infected mice had 40 to 80% positive at the start of the culture. We had previously shown that these macrophages rapidly lost the capacity to synthesize Ia, and became Ia-negative after 24 to 48 hours of culture¹⁰. The addition of a lymphokine-containing conditioned medium from immune T cells at the start of the culture resulted in the progressive expression of Ia

so that most of the macrophages were positive after 4 to 7 days regardless of the initial level of Ia-positive cells¹³. This expression of Ia was markedly inhibited by PGE₁ and PGE₂ in a dose-dependent manner with 50% inhibition (I₅₀) at 1.5×10^{-10} M and 8.6×10^{-10} M, respectively (Fig. 1). The dose of PGE required to inhibit Ia was dependent on the amount of T cell-conditioned medium. A fivefold increase in the amount (v/v) of conditioned medium relative to that used in the experiment reported in Fig. 1 did not change the per cent of macrophages that became Ia-positive but resulted in a 10-fold increase in the I₅₀ dose of PGE₁ or E₂. PGE at concentrations of 10^{-7} M or less had no effect on macrophage viability, morphology, adherence properties, or on the per cent expressing IgG Fc receptors, or H-2K glycoproteins (Table 1). Identical results were found testing resident macrophages or those from exudates induced with peptone or following *Listeria* infection. It appears that PGE inhibition is associated with suppression of Ia biosynthesis (unpublished observation). The time of addi-

Table 1 Effects of PGE₂ on expression of macrophage surface proteins

Treatment		% Of cells positive for:		
Lymphokine	PGE ₂	Ia	H-2K	Fc receptors
–	–	1	78	>98
+	–	37	83	>98
+	10^{-9} M	18	82	>98
+	10^{-7} M	5	82	>98
+	10^{-6} M	2	77	>98

Macrophages were prepared and cultured exactly as described in Fig. 1. Macrophages were cultured for 6 days with 2% conditioned media (lymphokine). Prostaglandins were added during the last 2 days of culture. Ia expression was assessed by immunofluorescence as in Fig. 1. H-2K expression was similarly determined using a monoclonal anti-H-2K^k antibody (clone line 11-4.1, from Dr L. Herzenberg), followed by fluorescein-labelled F(ab)₂ rabbit anti-mouse immunoglobulin. The cells were fixed after staining for H-2K. Fc receptors were detected by standard techniques of erythrocyte rosetting, using sheep red blood cells coated with rabbit IgG antibody.

tion of PGE to the cultured macrophages was important. Addition for only the first 2 days retarded the appearance of Ia-positive cells by 1 to 2 days but did not affect the ultimate level of Ia expression; addition at the time when most macrophages were Ia-positive resulted in loss of expression within 48 h.

As PGE is known to increase intracellular cyclic AMP levels in macrophages¹⁴, we tested the effects of cyclic AMP analogues

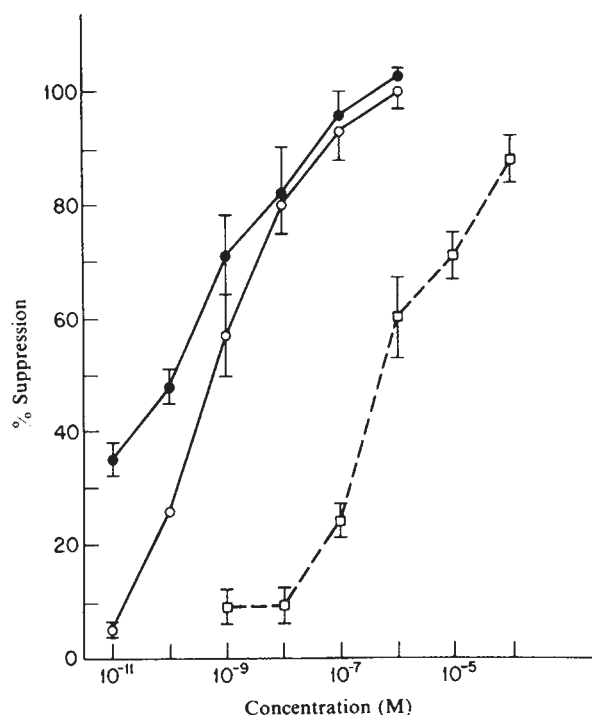


Fig. 1 Suppression of macrophage Ia expression by PGE₁ (●), PGE₂ (○), and dibutyryl cyclic AMP (□) in culture. Macrophages were obtained from the peptone-elicited peritoneal exudates (1 ml 10% proteose-peptone, 3 days before cells were collected) of adult A/St mice previously infected with 2×10^4 live *Listeria monocytogenes* organisms intraperitoneally²². The exudates were collected 3 to 4 days later by lavaging with 5 ml of Hank's balanced salt solution containing 0.6% g bovine serum albumin, 10 mM HEPES, and 10 units per ml heparin. Peritoneal exudate cells (2×10^5) were plated on glass coverslips placed in flat-bottom, Costar wells, in 1 ml of RPMI 1640 containing 5% fetal calf serum for 2.5 h at 37 °C, 5% CO₂; then non-adherent cells were washed off. The medium was replaced with RPMI 1640 containing 10% fetal calf serum, plus or minus 2% conditioned media from a *Listeria*-immune T-cell line¹³. This medium was washed off on day 3 or 4 and replaced with fresh RPMI with or without the drugs. Prostaglandins were the gift of Dr John Pike (Upjohn). They were stored as stock solution at 10 mg per ml of absolute ethanol at 4 °C, then diluted before each experiment with a 0.2 M phosphate buffer. On day 4 or 5, the macrophages were given a 24-h pulse of 5×10^6 heat-killed *Listeria* organisms and were then fixed in 1% paraformaldehyde and stained for surface Ia on day 5 or 6²². Cells on glass coverslips were exposed to a monoclonal anti-I-A^k antibody (clone 10-2.16) (5 µg per ml) mixed with a 1:5 dilution of decomplexed normal rabbit serum, at 4 °C for 15 min. Fluorescent labelling was carried out as described in ref. 22. The % of Ia-positive macrophages was determined by counting 200 cells with a fluorescent microscope. Macrophages cultured in RPMI alone were 0 to 5% Ia-positive; those cultured in T-cell lymphokine were 40 to 80% positive. Per cent suppression was calculated as:

$$\frac{\% \text{ Ia-positive control} - \% \text{ Ia-positive experimental}}{\% \text{ Ia-positive control} - \% \text{ Ia-positive background}} \times 100$$

Data represent the mean $\pm$ s.e.m. These results are similar to those from radiolabelled antibody binding.

or agonists. Figure 1 shows that dibutyryl cyclic AMP inhibited Ia expression with an I_{50} of 9×10^{-7} M. (In other experiments, isoprenaline inhibited in the dose range of 10^{-5} M to 10^{-6} M and theophylline in the range 10^{-3} to 10^{-4} M.)

The concentrations of PGE inhibiting Ia are well within the physiological range found in tissues, which varies from about 10^{-11} M in basal states¹⁵ to as high as 10^{-7} M in inflammatory sites¹⁶. The inhibitory effects of PGE on macrophage Ia could be shown *in vivo* under two circumstances (Table 2). First, intraperitoneal injection of PGE₁ inhibited the lymphokine-induced development of peritoneal exudates rich in Ia-positive macrophages¹¹. Second, injection of indomethacin not only augmented the response to lymphokines but increased the basal number of Ia-positive macrophages as well. Further indication of the role of PG *in vivo* comes from an earlier study in which we showed that repeated injection of resident peritoneal macrophages—the phagocyte population producing the highest basal level of PGE¹⁷—interfered with the lymphokine-induced development of Ia-positive macrophages¹². Additionally, there are reports that indomethacin stimulates monocytopenia resulting in increased numbers of peripheral blood monocytes and tissue macrophages¹⁸ and that PGE inhibits monocytopenia¹⁹. Taken together, these examples suggest that PGE-mediated inhibition of Ia expression may be one of several regulatory effects exerted by arachidonic acid metabolites on the macrophage lineage.

Another important aspect of PGE regulation of Ia expression is the possibility that other arachidonic acid metabolites can antagonize the effects of PGE. We, therefore, looked at the effect of TXB₂, a stable metabolite of thromboxane A₂. Thromboxane A₂, which is produced by activated macrophages²⁰, has been shown to lower intracellular cyclic AMP levels²¹. Although TXB₂ is known to be less active than thromboxane A₂ it exerted an effect in our system. Figure 2 indicates that TXB₂ was a surprisingly potent antagonist to the effects of PGE₂ but had a small and inconsistent effect on Ia expression in the absence of T-cell stimulation. PGE₁ had no effect, either *in vitro* or *in vivo*. We are now investigating products of the lipooxygenase pathway. In preliminary studies, we have tested three monohydroxyeicosatetraenoic acids (5-HETE, 11-HETE, and 15-HETE provided by Dr E. Goetzl). In the presence of lymphokines, these lipooxygenase metabolites increased the number of Ia-positive macrophages by 13 to 60%.

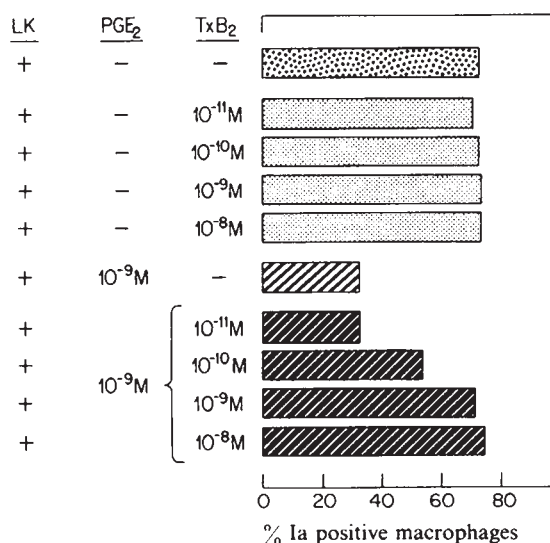


Fig. 2 TXB₂ and PGE₂ are antagonists. Macrophages were prepared and cultured as described in Fig. 1. On day 3, RPMI plus or minus 2% T-cell lymphokine (LK) was replaced with fresh RPMI containing PGE₂, TXB₂, or both. Cells were pulsed with heat-killed *Listeria* organisms on day 4, then fixed and stained for Ia on day 5.

Table 2 Effects of indomethacin and PGE₁ on peritoneal exudate macrophages

Experiment no.	Injected material	Macrophages per mouse ($\times 10^{-5}$)	Ia-positive (%)	Ia-positive macrophages per mouse ($\times 10^{-5}$)
1	Indomethacin, 50 μ g in 5% ethanol 5% Ethanol	10.0	5	0.5
		24.3	28	7.0
		19.3	8	1.5
2	Indomethacin, 50 μ g in 5% ethanol Lymphokine, 1:20 Indomethacin, 50 μ g, plus lymphokine, 1:20	5.5	9	0.5
		26.0	23	6.0
		11.6	38	4.4
		26.2	50	13.1
3	Lymphokine, 1:2 Lymphokine, 1:2, plus 15-methyl PGE ₁ , 1 μ g	8.6	4	0.3
		30.2	35	10.6
		26.0	15	3.9
4	15-Methyl PGE ₁ , 1 μ g PGE ₁ , 50 μ g Lymphokine, 1:2 Lymphokine, 1:2, plus 15-methyl PGE ₁ , 1 μ g Lymphokine, 1:2, plus PGE ₁ , 50 μ g	8.8	4	0.35
		24.8	9	2.2
		15.4	7	1.1
		35.0	45	15.8
		57.8	8	4.6
		47.5	5	2.4

A/St or B10.A adult mice were injected intraperitoneally with 0.5 ml of the materials every 12 h, six times, as previously reported¹¹. Peritoneal exudate cells were collected 12 h after the last injection, then plated on coverslips in RPMI for 2.5 h as described in Fig. 1. Non-adherent cells were washed off and counted, and the per cent adherent cells was calculated. Adherent cells were fixed and stained for Ia as described in Fig. 1.

Overall, our results suggest that prostaglandins can regulate the expression of Ia antigens on macrophages induced by lymphokines (Fig. 1, Table 1) as well as the T-cell-independent development of the basal levels of Ia-positive macrophages in tissues. Our results also suggest that the relative concentration of different arachidonic acid metabolites influences the level of Ia expression (Fig. 2). The finding that indomethacin increased Ia expression *in vivo* in the absence of exogenous lymphokine suggests either that PGE interferes with the spontaneous expression of Ia by some macrophages or that it influences the response to environmental stimuli. The basal level of Ia-positive macrophages, however, was not inhibited by injection of PGE (Table 1), implying that there may be some heterogeneity among macrophages with respect to PGE sensitivity. Finally, we suggest that the regulation of Ia expression by prostaglandins could be an important control mechanism in the inductive phase of immune responses. We have recently found that the loss of Ia by macrophages after exposure to PGE is associated with reduced antigen presentation to immune T cells (unpublished observations).

This work was supported by grants from the US NIH and the Council for Tobacco Research Inc.

Received 5 April; accepted 28 June 1982.

- Kueh, F. A. Jr & Egan, R. W. *Science* **210**, 978-984 (1980).
- Goldyne, M. E. & Stobo, J. D. *CRC crit. Rev. Immun.* **2**, 189-223 (1981).
- Stenson, W. F. & Parker, C. W. *J. Immun.* **125**, 1-5 (1980).
- Scott, W. A., Zrike, J. M. Hamill, A. L., Kempe, J. & Cohn, Z. A. *J. exp. Med.* **152**, 324-335 (1980).
- Rouzer, C. A., Scott, W. A., Kempe, J. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4279-4282 (1980).
- Kurland, J. I. & Bockman, R. J. *exp. Med.* **147**, 952-957 (1978).
- Leung, K. H. & Mihich, E. *Nature* **288**, 597-600 (1980).
- Webb, D. R. & Nowowiejski, I. *Cell. Immun.* **33**, 1-10 (1977).
- Unanue, E. R. *Adv. Immun.* **31**, 1-136 (1981).
- Beller, D. I. & Unanue, E. R. *J. Immun.* **126**, 263-269 (1981).
- Scher, M. G., Beller, D. I. & Unanue, E. R. *J. exp. Med.* **152**, 1684-1698 (1980).
- Snyder, D. S., Lu, C. Y. & Unanue, E. R. *J. Immun.* **128**, 1458-1465 (1982).
- Beller, D. I. & Ho, K. J. *Immun.* **129** (in the press).
- Bonta, I. L., Adolfo, M. J. P. & Parnham, M. J. *Prostaglandins* **22**, 95-103 (1981).
- Samuelsson, B. in *Advances in the Biosciences* (ed. Bergstrom, S.) 7-23 (Pergamon, Oxford, 1973).
- Higgs, G. A. & Salmon, J. A. *Prostaglandins* **17**, 737-746 (1979).
- Humes, J. L. *et al. J. Immun.* **124**, 2110-2116 (1980).
- Razin, E., Hayari, Y. & Globerson, A. *Prostaglandins Med.* **6**, 613-620 (1981).
- Pelus, L. M., Broxmeyer, H. E., Kurland, J. I. & Moore, M. A. S. *J. exp. Med.* **150**, 277-292 (1979).
- Brune, K., Glatt, M., Kalin, H. & Peskor, B. A. *Nature* **274**, 261-263 (1978).
- Moncada, S. & Vane, J. R. *New Engl. J. Med.* **300**, 1142-1147 (1979).
- Beller, D. I., Kiely, J.-M. & Unanue, E. R. *J. Immun.* **124**, 1426-1432 (1980).

Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line

M. W. McBurney, E. M. V. Jones-Villeneuve,
M. K. S. Edwards & P. J. Anderson

Departments of Medicine, Biology and Biochemistry,
University of Ottawa, Ottawa, Canada K1H 8M5

Pluripotent murine embryonal carcinoma cells can differentiate in culture into many tissue types similar to those normally found in early embryos¹ and may be useful in investigating some developmental events^{1,2}. Central to our understanding of embryonic development are explanations of cellular determination, that is, the commitment of early embryonic cells to form divergent cell types. Of relevance is recent work with the F9 line of embryonal carcinoma cells which suggests that certain extra-embryonic cell types are specifically formed following treatment of undifferentiated cells with drugs^{3,4} and the manipulation of culture conditions⁵. We report here that the P19 line of embryonal carcinoma cells⁶ may provide an analogous system in which drugs can be used to manipulate the formation of tissues which normally comprise the fetus. In the presence of dimethyl sulphoxide (DMSO) aggregates of P19 cells differentiate rapidly to form large amounts of cardiac and skeletal muscle but no neurones or glia. We have previously shown that in the presence of high concentrations of retinoic acid ($>5 \times 10^{-7}$ M), aggregates of these same cells develop into neuronal and glial tissues but not muscle⁷. Thus, drugs can be used to generate two quite different spectra of embryonic tissue types from the same population of embryonal carcinoma cells.

P19 is a euploid (40:XY) embryonal carcinoma cell line derived from a teratocarcinoma induced in C3H/He strain mice⁶. For the experiments described below, we used P19S1801A1, a ouabain-resistant and 6-thioguanine-resistant subclone of P19 isolated without mutagenesis. Suspensions of dispersed cells were plated onto bacterial-grade plastic surfaces to which cells do not adhere⁸. Cells adhere to each other to form small aggregates. These aggregates were cultured in suspension for 4-5 days in the presence or absence of DMSO.

They were then plated into tissue culture-grade plastic dishes.

In the absence of drug, the plated aggregates contained undifferentiated embryonal carcinoma cells along with small numbers of extra-embryonic endodermal cells⁷ (Fig. 1*a*). The presence of DMSO in the culture medium produced effects which became clear 1–2 days after plating, that is, 6–7 days after initiation of the experiment. In cultures exposed to 0.25% (v/v) DMSO, most plated aggregates contained embryonal carcinoma cells, rhythmically contracting muscle and fibroblast-like cells. At concentrations of 0.5, 0.75 and 1.0% DMSO, none of the plated aggregates contained embryonal carcinoma cells (identified by morphology), virtually all contained areas of rhythmically contracting muscle, and all contained cells with fibroblast-like morphology (Fig. 1*c*). By 10–12 days the amount of contracting muscle had increased (Fig. 1*d, e*). Also at this time many of the DMSO-treated aggregates developed areas of bipolar myoblasts which fused into myotubes (Fig. 1*f*). These myotubes were usually non-contractile but often developed spontaneous twitching activity by 14 days.

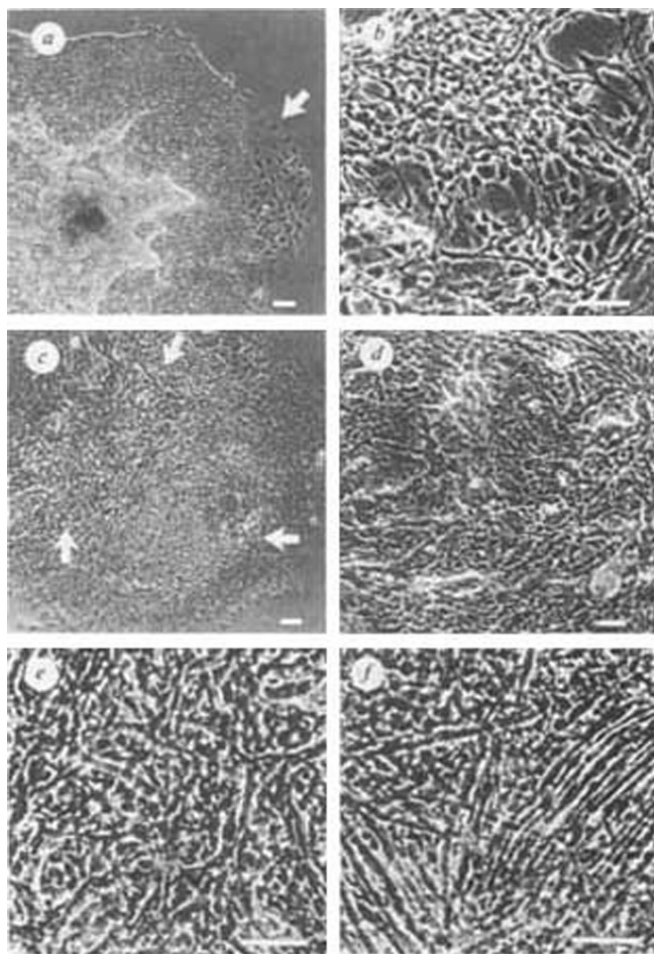


Fig. 1 Phase contrast photomicrographs of live teratocarcinoma cells. The conditions for cell culture^{7,12} and aggregation⁸ have been described previously. Cell aggregates were formed from a stock culture of P19S1801A1 cells and parallel cultures were carried in: *a*, normal medium (α -medium plus 10% fetal bovine serum¹²); *b*, in medium containing 5×10^{-7} M retinoic acid; and *c–f*, in medium containing 0.5% (v/v) DMSO. Aggregates were cultured in suspension in bacterial-grade Petri dishes for 5 days before being plated on to tissue culture-grade plastic surfaces. Photographs were taken 2 days (*a–c*) or 9 days (*d–f*) later. *a*, Untreated aggregates contain embryonal carcinoma and a few extra-embryonic endoderm cells (arrow). *b*, Aggregates of cells which had been cultured in the presence of 5×10^{-7} M retinoic acid contain neurones and astrocyte glial cells⁷. *c–f*, Aggregates cultured continuously in the presence of 0.5% DMSO contain small areas of rhythmically contracting cardiac muscle (arrows in *c*) which become more extensive with time (*d*). At higher magnification are: *e*, areas of rhythmically contracting mononucleate cardiac muscle and, *f*, multinucleate skeletal muscle. Scale bars, 200 μ m.

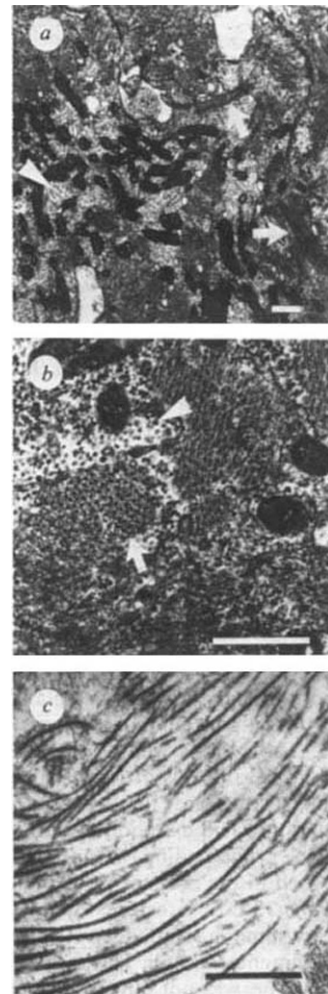


Fig. 2 Electron micrographs of some of the tissues formed in DMSO-treated cultures. *a*, A section through cardiac muscle shows bundles of thick and thin filaments in longitudinal section (arrows), glycogen granules (arrowhead) and numerous mitochondria. *b*, A section through a multinucleate myotube shows the thick and thin filaments in cross section (arrow) and glycogen (arrowhead). Many of the non-muscle cells in these cultures secrete collagen (*c*) which was often seen forming part of an intercellular matrix. Scale bars, 0.5 μ m.

Electron microscopy of the cells in DMSO-treated cultures indicated that the rhythmically contracting cardiac muscle cells contained glycogen granules, large numbers of mitochondria, and numerous areas of thick and thin filaments which were not organized into mature myofibrils (Fig. 2*a*). The multinucleate skeletal muscle cells were similar in appearance (Fig. 2*b*). Thus both muscle types seemed to be immature. Many of the non-muscle cells had abundant rough endoplasmic reticulum and some were surrounded by extracellular matrix which included collagen fibres (Fig. 2*c*).

The DMSO-treated aggregates of P10S1801A1 cells developed muscle but neither neurones nor glia. Treatment of the same cells with retinoic acid resulted in the development of neurones (Fig. 1*b*), glial cells, fibroblast-like cells, but no muscle. Cultures exposed to both retinoic acid (5×10^{-7} M) and DMSO (0.5 or 1.0%) developed as if exposed only to retinoic acid, that is, neurones and glia but no muscle were formed.

Differentiated cultures contained more actin than did untreated cultures (Table 1). Much of the actin in DMSO-treated cultures was α -actin, the type present only in skeletal and cardiac muscle cells⁹. Muscle-specific myosin was also detected in both cardiac and skeletal muscle by immunofluorescence using monoclonal antibodies directed against muscle myosin. About 15% of all cells were muscle myosin positive in these cultures by 8 days but none were detected in untreated or in retinoic acid-treated cultures.

When 10 μ M adrenaline was added to DMSO-treated cultures, the cardiac muscle responded by a 2–2.5-fold increase in contraction frequency and some previously quiescent areas of the culture were stimulated into rhythmic activity. Therefore β -adrenergic receptors were present. Such receptors are apparently acquired by cardiac muscle after the acquisition of spontaneous contractility¹⁰.

DMSO was not demonstrably cytotoxic to the P19S1801A1 cells at concentrations effective in differentiation experiments. Figure 3 shows that the efficiency of colony formation was unaffected by DMSO at concentrations up to 1.0%. Virtually all colonies formed in DMSO contained only embryonal carcinoma cells. In other experiments, monolayers of cells were cultured for 20 days in 1% DMSO without change in growth rate or morphology. At the end of this 20-day period, the DMSO-treated cells were aggregated in the presence or absence of DMSO (0.5%). Those aggregates formed in the absence of DMSO did not differentiate while those cultured in the drug formed muscle and fibroblasts in the usual way. Thus, it seems that the DMSO had no effect on the P19S1801A1 cells cultured as monolayers and that both the drug and cell aggregation are necessary for muscle differentiation. DMSO could be removed after 2–3 days but cardiac muscle still developed at 6–7 days as in the continuous presence of the drug.

The effects of DMSO described above were observed not only on P19S1801A1 cells, but also on the parental P19 cells and on all of the subclones from this line which were tested. However, DMSO had no effect on the differentiation of the embryonal carcinoma cell lines F9¹¹, OC15S1¹² and C86S1¹² whereas some clones of P10 cells¹³ appear to form an excess of neurones in the presence of DMSO (G. D. Paterno and M.W.M., unpublished). Variation has also been observed in the response of different embryonal carcinoma lines to retinoic acid^{3,7} and to aggregation in the absence of drugs^{1,2}.

DMSO is an inducer of Friend cell differentiation¹⁴ as are 6-thioguanine¹⁵, butyrate¹⁶ and ouabain¹⁷. The effects reported above for DMSO have also been observed with non-toxic concentrations of 6-thioguanine and butyrate but not with ouabain. Another Friend cell inducer, hexamethylene bisacetamide (HMBA), has previously been shown to influence the differentiation of some other embryonal carcinoma cell lines^{18,19} but we have not tested this compound with P19 cells.

Many papers have reported the formation of a limited range of cell types following spontaneous or induced differentiation of lines of teratocarcinomas and of embryonal carcinoma cells^{20–23}, but it is not clear whether this is the result of differential selection or of the occurrence of a limited number of determinative events. We think it is unlikely that differential selection can account for our observations because: (1) the cells

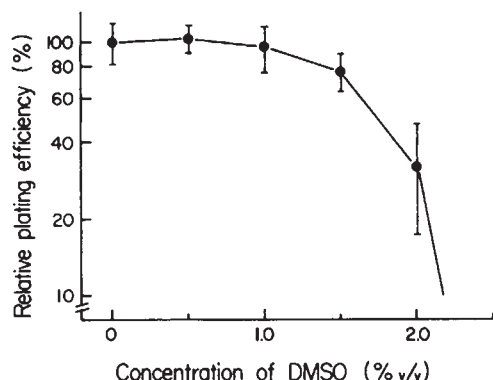


Fig. 3 The efficiency of colony formation of P19S1801A1 embryonal carcinoma cells in the presence of DMSO indicates the absence of toxicity at concentrations of less than 1% (v/v). About 200 cells were introduced into replicate 60-mm diameter dishes containing various concentrations of DMSO dissolved in α -medium supplemented with 10% fetal bovine serum and 10^{-4} M β -mercaptoethanol. Incubation was for 8 days at 37°C. All colonies consisted of cells with embryonal carcinoma morphology. Those colonies formed in 1.5% DMSO were smaller than control colonies, indicating that at these concentrations, the rate of cell proliferation was decreased.

Table 1 Presence of α -actin in DMSO-treated cultures

Treatment	Total actin*	α -actin*	% Muscle actin
Untreated, day 7	1.91 $\pm$ 0.07	0.16 $\pm$ 0.02	8.4
Retinoic acid, day 7	2.31 $\pm$ 0.03	0.30 $\pm$ 0.02	13.0
DMSO, day 7	2.62 $\pm$ 0.15	0.52 $\pm$ 0.01	19.8
DMSO, day 11	3.05 $\pm$ 0.14	0.68 $\pm$ 0.02	22.3

Aggregated cultures were prepared as described in Fig. 1 legend. Two days after plating (day 7) or 6 days after plating (day 11) the cultures were collected for analysis. The DMSO-treated day 7 culture contained rhythmically contracting but not multinucleate muscle. By day 11 both muscle types were present. Protein and peptide isolation was carried out as previously described²⁴ except for the use of trypsin instead of chymotrypsin for peptide generation. The actin contents were calculated by measuring the amounts of material which co-purified during electrophoresis at pH 6.5, 2.1 and 3.5 with tryptic peptides generated from muscle actin. Total actin values were calculated from the amounts of radioactive material co-migrating with the two chemically modified peptides CmCys-Asp-Ile-Asp-Ile-Arg and CmCys-Phe. All known actins contain peptides which should co-purify with these two. α -Actin values were calculated from the amount of material co-purifying with an 18-residue CmCys-containing peptide generated from the N-terminal region of α -actin. Actins from other tissues differ from α -actin in this region⁹ so should not co-purify with this peptide. Total actin is higher in differentiated than in undifferentiated cultures and there is more muscle-specific α -actin in DMSO-treated cultures. The 8–13% of muscle actin present in untreated and in retinoic acid-treated cultures may represent a background of radioactive label derived from non- α -actin peptides which co-purify with the legitimate peptide.

* mg actin per 100 mg total protein.

did not differentiate into embryonic cell types in the absence of drugs, (2) the cell types formed in DMSO-treated cultures were substantially different from those formed by the same cells in parallel cultures exposed to retinoic acid, (3) neither drug appeared to be toxic, (4) all subclones responded to both drugs, (5) the drugs were effective even when cultures were exposed to them for 48 h at the beginning of an experiment, and (6) DMSO did not inhibit the formation of neurones in cultures exposed to both retinoic acid and DMSO. The simplest interpretation of these data seems to be that each drug acts by 'inducing' uncommitted embryonal carcinoma cells to differentiate along a limited number of developmental avenues. If the drugs act by bringing about intracellular changes which mimic the results of certain embryonic decisions, it may be possible to use the drugs to identify parts of the cellular decision-making apparatus.

This work was supported by grants from the NCI of Canada and the MRC of Canada. We thank Irwin Schweitzer and Kem Rogers for their help. Antibody to glial fibrillar protein was the gift of Dr V. Kalnins and antibody to muscle-specific myosin was provided by Drs D. Morgenstern, D. A. Fischman and P. Merrifield.

Received 10 January; accepted 15 June 1982.

- Graham, C. F. in *Concepts in Mammalian Embryogenesis* (ed. Sherman, M. I.) 313–394 (MIT Press, Cambridge, 1977).
- Martin, G. R. *Science* **209**, 768–775 (1980).
- Strickland, S. & Mahdavi, V. *Cell* **15**, 393–403 (1978).
- Strickland, S., Smith, K. K. & Marotti, K. R. *Cell* **12**, 347–355 (1980).
- Hogan, B. L. M., Taylor, A. & Adamson, E. *Nature* **291**, 235–237 (1981).
- McBurney, M. W. & Rogers, B. J. *Dev Biol.* **89**, 503–508 (1982).
- Jones-Villeneuve, E. M. V., McBurney, M. W., Rogers, K. A. & Kalnins, V. I. *J. Cell Biol.* (in the press.)
- Martin, G. R. & Evans, M. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1441–1445 (1975).
- Vandekerckhove, J. & Weber, K. *Eur. J. Biochem.* **113**, 595–603 (1981).
- Lipshultz, S., Shanfeld, J. & Chacko, S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 288–292 (1981).
- Bernstine, E. G., Hooper, M. L., Grandchamp, S. & Ephrussi, B. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3899–3903 (1973).
- McBurney, M. W. *J. cell. Physiol.* **89**, 441–456 (1976).
- McBurney, M. W. & Strutt, B. J. *Cell* **21**, 357–364 (1980).
- Friend, C., Scher, W., Holland, J. G. & Sato, T. *Proc. natn. Acad. Sci. U.S.A.* **68**, 378–383 (1971).
- Gusella, J. F. & Housman, D. *Cell* **8**, 263–269 (1976).
- Leder, A. & Leder, P. *Cell* **5**, 319–322 (1975).
- Bernstein, A., Hunt, D. M., Crichtley, V. & Mak, T. W. *Cell* **9**, 375–381 (1976).
- Jakob, H., Dubois, P., Eisen, H. & Jacob, F. *C. r. heb. Séanc. Acad. Sci., Paris* **286D**, 109–111 (1978).
- Speers, W. C., Birdwell, C. R. & Dixon, F. J. *Am. J. Path.* **97**, 563–584 (1979).
- Gearhart, J. G. & Mintz, B. *Cell* **6**, 61–66 (1975).
- Vandenberg, S. R., Herman, M. M., Ludwin, S. K. & Bignami, A. *Am. J. Path.* **79**, 147–168 (1975).
- Pfeiffer, et al. *J. Cell Biol.* **88**, 57–66 (1981).
- Darmon, M., Bottenstein, J. & Sato, G. *Dev Biol.* **85**, 463–473 (1981).
- Anderson, P. J. *Biochem. J.* **179**, 425–430 (1979).

A new fucosyl antigen expressed on colon adenocarcinoma and embryonal carcinoma cells

Teruo Miyauchi*, Suguru Yonezawa†, Tsukasa Takamura‡, Taku Chiba‡, Setsuzo Tejima‡, Masayuki Ozawa*, Eiichi Sato† & Takashi Muramatsu*

Departments of *Biochemistry and †Pathology, Kagoshima University School of Medicine, 1208-1 Usukicho, Kagoshima 890, Japan

‡ Department of Hygienic Chemistry, Faculty of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabedori, Mizuhoku, Nagoya 467, Japan

Carbohydrate antigens on the surface of mammalian cells have recently gained renewed interest because some are specifically expressed at certain stages of cellular differentiation¹⁻⁴. Most of the useful antibodies detecting such carbohydrate antigens have been monoclonal antibodies produced by the hybridoma technique using whole cells as the immunogen¹⁻³ or the antibodies in the sera of some of human patients⁴. Here we report that an antiserum raised against an unusual carbohydrate linkage prepared by organic synthesis can preferentially react with certain tumour cells. Thus, an antiserum against the Fucal→3Gal linkage reacted with human colon adenocarcinoma cells and murine teratocarcinoma cells but reacted only with severely restricted regions in normal tissues.

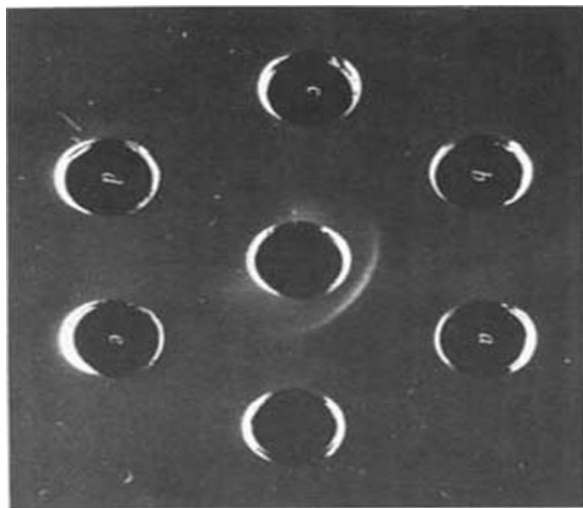


Fig. 1 Ouchterlony double-diffusion analysis on the specificity of the rabbit antiserum against 3'-fucosyllactose-PIP-BSA. The antigen (0.4 mg) was mixed with Freund's complete adjuvant and injected into the footpads of a rabbit. After 3 and 5 weeks, the rabbit received two booster injections using half the amount of antigen. The rabbit was bled 10 days after the last immunization. The double-diffusion analysis was performed in 1% agar gel containing 2% Triton X-100, 0.15 M NaCl, 50 $\mu\text{g ml}^{-1}$ of phenyl-methylsulphonyl fluoride and 0.05% sodium azide in 0.01 M Tris-HCl buffer, pH 7.6, on a slide glass. The central well contained the antiserum and the peripheral wells contained 20 μg each of the following materials: a, 3'-fucosyllactose-PIP-BSA, b, 4'-fucosyllactose-PIP-BSA, c, 6'-fucosyllactose-PIP-BSA, d, lactose-PIP-BSA, e, BSA. Preparation of 4'-fucosyllactose and 6'-fucosyllactose has been described elsewhere¹². Carbohydrate contents of the PIP-BSA derivatives were determined by the phenol- H_2SO_4 reaction¹³, and were found to be 20 mols per mol of BSA in 3'-fucosyllactose-PIP-BSA, 25 mols per mol of BSA in the 4'-fucosyl compound and 23 mols per mol of BSA in the 6'-fucosyl compound.

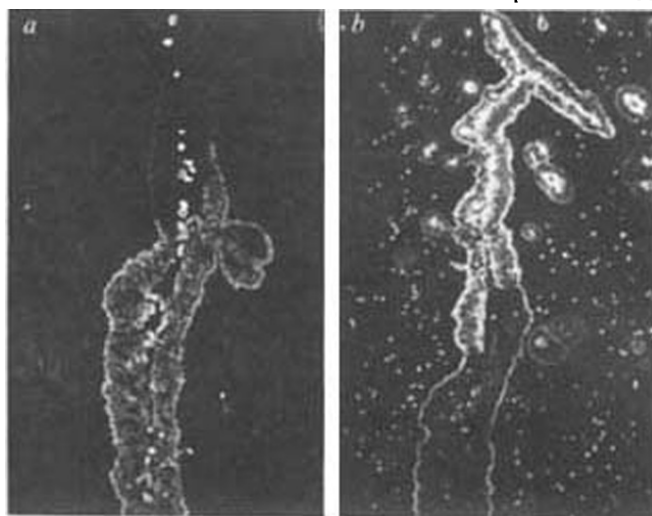


Fig. 2 Reaction of the serial tissue sections from a mouse uterus with the anti-Fucal→3Gal serum or with UEA-1. The organs were taken from a 3-month-old female 129/SV mouse, quick frozen and sectioned in a cryostat. The section was fixed for 1 min with acetone. a, Immunofluorescence staining with the antiserum showing the presence of the antigen in the squamous epithelium of the uterine cervix. The section was incubated with the antiserum diluted 20-fold with phosphate-buffered saline and then with FITC-conjugated F(ab')_2 fragment of goat anti-rabbit IgG (Cappel), and was observed by a fluorescence microscope using epillumination (Olympus, Model BH-RFL-LB). $\times 80$. Very similar results were also obtained using BALB/c, ICR and ddY mice. No staining was observed in control experiments where rabbit non-immune serum, rabbit anti-BSA or rabbit anti-lactose-PIP-BSA was used instead of the specific antiserum. b, Staining with FITC-UEA-1. The tissue sections were incubated with 25 μl of FITC-UEA-1 at 1 mg ml^{-1} . $\times 80$.

3'-Fucosyllactose (Fucal→3Galβ1→4Glc) was synthesized as described previously⁵, converted to the *p*-isothiocyanate-phenethylamine (PIP) derivative and coupled with bovine serum albumin (BSA) according to Smith *et al.*⁶. The product (3'-fucosyllactose-PIP-BSA), which contained 20 mols of 3'-fucosyllactose per mol of BSA, was injected into a rabbit. The resulting antiserum formed a sharp precipitation line with the antigen, but not with BSA, lactose-PIP-BSA, 4'-fucosyllactose-PIP-BSA nor with 6'-fucosyllactose-PIP-BSA (Fig. 1). Furthermore, these BSA derivatives without the Fucal→3Gal linkage were unreactive to the antiserum even in a binding assay, in which Sepharose coupled with the derivatives were stained with the antiserum by indirect immunofluorescence technique used in Fig. 2.

The reactivity of the antiserum to normal cells was also studied. Frozen sections of organs of adult mice were incubated with the diluted antiserum, and then fluorescein isothiocyanate (FITC)-labelled F(ab')_2 fragment of goat anti-rabbit IgG. We found that the antiserum stained squamous epithelium of the uterine cervix (Fig. 2a), ciliated border of the oviduct, epithelium of the bronchus, ependyma lining the ventricles of the brain and spermatocytes of the testis. The following organs were negative for the antigen: small intestine, proximal and distal colons, liver, spleen, kidney, epididymis, ovary and endometrium. The extremely restricted distribution of the antigen in the mouse suggests that the antiserum does not cross-react with H antigen whose determinant is Fucal→2Gal linkage⁷ or with SSEA-1 whose determinant involves Fucal→3GlcNAc linkage^{2,3}. SSEA-1 is known to be present in the epididymis and in the endometrium but not in squamous epithelium of the uterine cervix. It is also present in a portion of epithelial cells of the oviduct but not in ciliated border of the oviduct⁸. H antigen detected by staining with FITC-conjugated *Ulex europaeus* agglutinin 1 (UEA-1) was expressed in a variety of the adult tissues including endometrium of the uterus (Fig.

2b), small intestine, proximal and distal colons. Furthermore, the present antiserum did not stain erythrocytes of a number of human subjects whose blood types included H and Le^a. The failure to react with Le^a erythrocytes excluded the possibility that the antiserum cross-reacts with Fuc α 1 $\rightarrow$ 4GlcNAc linkage⁷. From all these results, we propose that the antiserum specifically recognized Fuc α 1 $\rightarrow$ 3Gal linkage. The antigen recognized by the antiserum was termed as FG3 antigen.

The antigen was detected in human colonic adenocarcinoma (Fig. 3a). The staining reaction was inhibited by 0.2 M 3'-fucosyllactose, but not by 0.2 M lactose, 0.2 M fucose, nor by 10 mg ml⁻¹ BSA. So far frozen sections of 17 human colonic adenocarcinoma have been examined. Of these, seven cases expressed the antigen in intracytoplasmic granules (Fig. 3a) and five cases expressed it along the cell apex. As the antigen was present not only on the luminal surface of the carcinoma but also in the invasive nests of the carcinoma in deep layers of intestinal wall (Fig. 3a), we considered that the antigen was produced by the carcinomas themselves. Only tumour cells

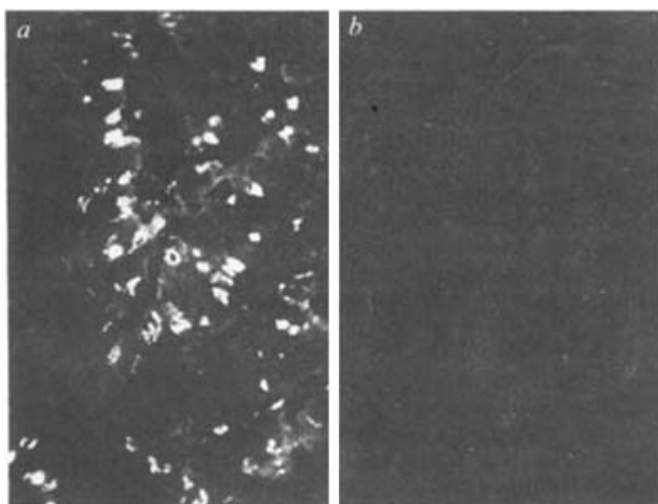


Fig. 3 Immunofluorescent staining of a human colonic adenocarcinoma using the anti-Fuc α 1 $\rightarrow$ 3Gal serum. The tissue section was prepared from surgically removed specimen of a human colon adenocarcinoma of 44-year-old female Japanese patient (Blood group B type). Staining was performed as described in the legend of Fig. 2. *a*, Cancerous lesion. $\times 160$. *b*, Non-neoplastic mucosa of this patient. $\times 80$.

within the colonic tissue section were positive for the antigen (Fig. 3b). Antiserum against lactose-PIP-BSA did not react with the tumour region. FG3 antigen was undetectable in tissue sections of normal human duodenum, liver, gallbladder, pancreas, lung, thyroid, thymus, lymph node, muscle, connective tissue and blood vessels. Although further studies are needed to examine the distribution of the antigen in other normal tissues and also in other neoplastic conditions, the tumour-specific expression of the antigen in human colon is of significant interest.

As stem cells of teratocarcinoma are known to express unusual fucosyl glycopeptides^{9,10}, we examined the expression of FG3 antigen in a stem cell line, F9 and found it (Fig. 4). The antigen was also detected on stem cells of teratocarcinoma OTT6050 and on another stem cell line, STT-M. Therefore, some embryonic antigens expressed on teratocarcinoma stem cells may contain this unusual fucosyl linkage.

However, the Fuc α 1 $\rightarrow$ 3Gal linkage is not found in glycoproteins and glycolipids isolated from diverse sources, but its presence has been reported in mixtures of higher homologues of human milk oligosaccharides¹¹. The present results suggest

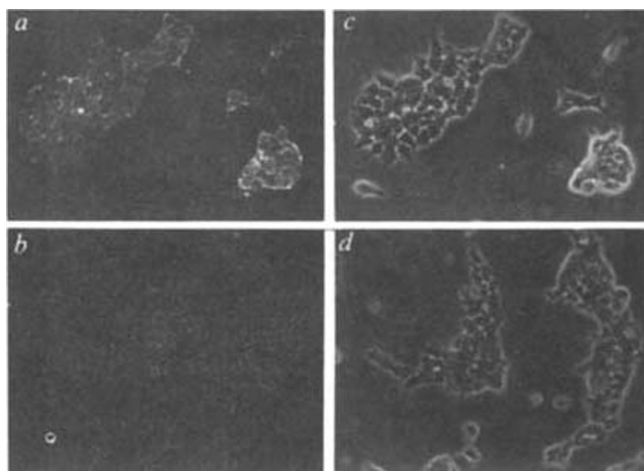


Fig. 4 Immunofluorescent staining of F9 embryonal carcinoma cells using the anti-Fuc α 1 $\rightarrow$ 3Gal serum. F9 cells were cultured on a glass coverslip in Dulbecco's modified minimum essential medium containing 15% fetal calf serum. Cells on the coverslip were stained as described in the Fig. 2 legend, except that acetone treatment was omitted. *a*, Specific staining. $\times 120$. *b*, Control staining using non-immune serum. $\times 120$. *c*, Phase contrast microscopic observation of the cells shown in *a*. $\times 120$. *d*, Phase contrast microscopic observation of the cells shown in *b*. $\times 120$.

that the unusual fucosyl linkage is expressed in certain cancerous and embryonic tissues. It is possible that such carbohydrate linkages are involved in intercellular communication during early embryonic life but are expressed aberrantly in certain cancerous cells.

We thank Miss Kumiko Sato for secretarial assistance.

Received 8 March; accepted 14 June 1982.

1. Stern, P. L. *et al. Cell* **14**, 775-783 (1978).
2. Solter, D. & Knowles, B. B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5565-5569 (1978).
3. Gooi, H. C. *et al. Nature* **292**, 156-158 (1981).
4. Kapadia, A., Feizi, T. & Evans, M. J. *Expl Cell Res.* **131**, 185-195 (1980).
5. Takamura, T., Chiba, T. & Tejima, S. *Chem. Pharm. Bull.* **29**, 1076-1082 (1981).
6. Smith, D. F., Zopf, D. A. & Ginsburg, V. *Meth. Enzym.* **50**, 169-177 (1978).
7. Watkins, W. M. *Science* **152**, 172-181 (1966).
8. Fox, N., Damjanov, I., Martinez-Hernandez, A., Knowles, B. B. & Solter, D. *Dev. Biol.* **83**, 391-398 (1981).
9. Muramatsu, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2315-2319 (1978).
10. Muramatsu, T., Gachelin, G., Dammonneville, M., Delarbre, C. & Jacob, F. *Cell* **18**, 183-191 (1978).
11. Yamashita, K., Tachibana, Y. & Kobata, A. *Archs Biochem. Biophys.* **174**, 582-591 (1976).
12. Chiba, T., Nishio, M., Takamura, T., Matsuda, H. & Tejima, S. *Proc. 3rd Symp. Carbohydrate Chemistry, Japan*, 90-91 (Japan Society Carbohydrate Chemists, 1980).
13. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. *Analyt. Chem.* **28**, 350-356 (1956).

In vivo surveillance of tumorigenic cells transformed in vitro

John Leslie Collins, Paul Q. Patek & Melvin Cohn

Salk Institute, PO Box 85800, San Diego, California 92138, USA

Any theory of surveillance against cancer requires that cells susceptible to host protective mechanisms exist as intermediates on the pathway from normal to cancer¹. The failure to demonstrate a significant frequency of such intermediates as a result of chemical carcinogenesis has cast serious doubt on the validity of the surveillance hypothesis². Here we report the conditions in which such intermediates can be identified as the major class of transformed cells resulting from *in vitro* chemical carcinogenesis of a cloned fibroblastic cell line.

We have previously reported the isolation and identification of cells that are intermediates between normal and cancerous cells³. We treated cloned cell lines (N-type) *in vitro* with

Table 1 Classification of cells by their tumorigenic potential

Type	Growth as a tumor in:	
	ATxFL mice	normal mice
N	—	—
I	+	—
C	+	+

chemical carcinogens and selected transformants by their anchorage-independent growth in agarose (Table 1). These transformants were then tested for their ability to grow as tumours in normal and surveillance-deficient mice; thymectomized, irradiated, and fetal liver restored (ATxFL). We found: (1) a correlation of 0.9 between the ability of these cells to grow in agarose and as tumours in ATxFL mice; (2) two classes of transformants, those which were tumorigenic in ATxFL mice but not in normal mice (I-type) and those which were tumorigenic in both (C-type); and (3) 20% of the chemically transformed cells were of the I-type and 80% of the C-type. The existence of I-type cells shows that a surveillance mechanism operates in normal mice to prevent the growth of some transformed cells and this surveillance mechanism is deficient in ATxFL mice.

The principle of the experiment described here is to treat N-cells *in vitro* with a chemical carcinogen and select transformants, not by their ability to grow in agarose, but by their ability to grow as tumours in ATxFL and normal mice. By testing the tumorigenic potential of carcinogen-treated N-cells in ATxFL and normal mice we were able to estimate the frequency with which transformation results in the I or C phenotype. Further, we demonstrated that transformants of N-cells selected for their tumorigenicity were anchorage-

independent, the converse of our previous experiment³ where cells selected for anchorage independence were tumorigenic. Thus, starting with this cloned non-tumorigenic, anchorage-dependent cell line, not only does selection for anchorage independence result in tumorigenicity but selection for tumorigenicity results in anchorage independence.

Benzo(a)pyrene epoxide (BPE) dissolved in tetrahydrofuran (THF) was added to culture dishes containing 10⁶ cloned N-cells in media without serum. The final concentration of BPE was 20 ng ml⁻¹, the medium contained 1% THF. Serum was added after 1 h of exposure to BPE; after 1 week the cells were collected, counted, and 10⁶ cells injected into ATxFL mice (experiment 1, Table 2) or ATxFL and normal mice (experiments 2 and 3, Table 2). As a control, an equal number of N-cells treated only with medium containing 1% THF were injected into mice.

Two points emerge from the results shown in Table 2. First, BPE dissolved in THF is much more effective than THF alone in transforming N-cells to grow as tumours; 46/57 dishes treated with BPE plus THF resulted in tumours compared with only 2/20 in the THF-treated dishes. Assuming that the number of transformants per dish follows a Poisson distribution, then the frequency of transformation by BPE plus THF is 40 times that of THF alone. Second, N-cells transformed by BPE can be divided into two classes based on their tumorigenicity, I-type growing as tumours in ATxFL but not normal mice and C-type growing as tumours in both ATxFL and normal mice. In the two experiments where the I or C phenotype of the transformants in 23 dishes was determined, 15 were I-type and 8 C-type. From the distribution of these data, I-type transformants appear at least three times more frequently than C-type.

The excised tumour cells were re-established in culture and then analysed for their ability to grow anchorage-independently. All of them expressed anchorage-independent growth (colony forming efficiencies³ between 1 and 30%) relative to the parental anchorage-dependent N-line (colony forming efficiency <0.01%) from which they were derived. The plating efficiency in agarose was independent of whether the tumour cells were derived from ATxFL or normal mice or whether tumours resulted from the injection of BPE-treated N-cells or from the few tumours formed by injecting control THF-treated N-cells. There is no correlation between the colony-forming efficiency of a transformant and its I or C phenotype.

The fact that all tumours, when re-established in culture, are anchorage-independent is complementary to our previous observation that anchorage-independent fibroblasts are tumorigenic³. Thus, for fibroblastic cell lines, tumorigenicity and anchorage independence are coordinately expressed whether the selection pressure used to isolate them is anchorage independence or tumorigenicity.

The major difference between the transformants selected by their ability to grow as tumours and those selected by their ability to grow in agarose is the frequency with which the transformations result in the I or C phenotype. By using transformants selected for their expression of anchorage independence, we previously estimated the frequency of transformation resulting in a cell of the I-type to be ~20% (ref. 3). In the experiments reported here the frequency of I-type transformants is much greater than 50%.

There are two possible explanations for the differences in the proportion of I- and C-type transformants in these two experiments. First, the *in vitro* selection for anchorage independence selects against anchorage-dependent transformants which grow as I-lines in mice. This means that there are both anchorage-independent and dependent transformants which grow as I-type cells in mice. When the transformants are tested directly in mice, both anchorage-independent and dependent transformants are revealed as I-type. When transformants are first selected for anchorage independence, then analysed in mice, only those which express anchorage independence are revealed. This reduces the total number of cells which express the I phenotype. It follows that if, in addition, there were

Table 2 The tumorigenic potential of N-cells treated *in vitro* with BPE

Treatment	No. of dishes containing cells which were:		Tumorigenic in ATxFL	Tumorigenic only in ATxFL mice (I-type)	Tumorigenic in ATxFL and normal mice (C-type)
	Non- tumorigenic				
Expt 1					
THF control	6		1		
BPE	1		23		
Expt 2					
THF control	2		0		0
BPE	1		7		1
Expt 3					
THF control	10		1		0
BPE	9		8		7

The N-cell line, B/C-N 7.1, was grown in Dulbecco's modified Eagle's minimum essential medium (DMEM) plus 10% fetal calf serum (FCS). Cells (10⁶ per 60 mm dish) in DMEM were treated with benzo(a)pyrene-*trans*-7,8-dihydrodiol-9,10-epoxide (BPE; supplied by NCI) dissolved in THF. The final concentration of BPE was 20 ng ml⁻¹ in 1% THF. The control dishes were treated with DMEM containing 1% THF. One week later the cells were removed from the dishes with trypsin, washed, and injected into two normal mice and two mice which had been thymectomized as adults, lethally irradiated (750 rad) and restored with 10⁷ syngeneic liver cells from 14–15-day-old-fetuses (ATxFL). Animals were examined weekly for tumours, which appeared between 45 and 120 days after injection. Tumours were excised when they reached approximate volumes of 1–2 cm³, mechanically dissociated, and re-established in *in vitro* culture. In all experiments tumours were derived from the injected parental N-cell populations and not from *in situ* transformation of host cells as shown by the fact that all the excised tumours were hypoxanthine-guanine phosphoribosyltransferase-negative, a marker introduced into the parental N-cells.

anchorage-dependent and independent transformants which were capable of growing as C-lines *in vivo*, the proportion of the I-type relative to the C-type would be the same in these two experiments. As the proportion of I-type varies, it seems that most transformants which express the C phenotype *in vivo* are anchorage-independent and most anchorage-dependent transformants are of the I phenotype.

Because no tumours were anchorage-dependent, any anchorage-dependent transformants must be at a stage on the pathway to cancer that allows the *in vivo* selection of anchorage-independent variants at a high frequency.

Second, the proportion of I- and C-type cells depends on the chemical carcinogen used to transform the N-line. When transformants were selected for anchorage independence before testing in mice, they were induced with various chemical carcinogens. All of the transformants tested directly in mice were induced with BPE. If BPE influences the tumorigenic phenotype of the transformants, then it must induce more I-type transformants than several other carcinogens that were analysed collectively.

The existence of the I phenotype shows that at least some transforming events result in cells which are sensitive to a tumour surveillance mechanism operating in normal but not ATxFL mice. ATxFL mice are severely depressed with respect to both immunological (cell-mediated and humoral immunity) and non-immunological (natural killer (NK) and natural cytotoxic (NC) cells) host protective mechanisms. Although it is possible that one (or both) of these host protective mechanisms is responsible for the destruction of I-lines by normal mice, we cannot ascribe the surveillance measured here to either mechanism. Regardless of the mechanisms of surveillance, the frequency with which transformation by a chemical carcinogen results in the I phenotype is at least three times that of the C phenotype. This is important because the efficacy of any surveillance mechanism is a function of the proportion of newly transformed cells which are susceptible to surveillance.

We acknowledge the contribution of the Armand Hammer Cancer Workshops and the technical assistance of Ms Scotti Brauer. This work was supported by NCI grant CA 19754-07.

Received 26 April; accepted 15 July 1982.

1. Burnet, M. *Immunological Surveillance* (Pergamon, London, 1970).
2. Prehn, R. T. *Am. J. Path.* **77**, 119-122 (1974).
3. Patek, P. Q., Collins, J. L. & Cohn, M. *Nature* **276**, 510-511 (1978).

A transforming gene present in human sarcoma cell lines

C. J. Marshall, A. Hall & R. A. Weiss

Chester Beatty Laboratories, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

Morphological transformation of NIH/3T3 cells by transfection with DNA has been used to identify transforming sequences in human tumours^{1,2}. Transforming activity has been reported for DNAs isolated from bladder, mammary, colon and lung carcinomas, neuroblastoma, lymphoid and myeloid tumours¹⁻⁶. Each of these tissues seems to contain different transforming sequences except for the colon and lung tumours where the same sequence seems to be involved⁵. We now report that in two different human sarcoma cell lines, a fibrosarcoma and an embryonal rhabdomyosarcoma, the DNAs have transforming activity. The transforming gene is the same in both sarcomas but differs from the activated sequences detected in other tumours. We have also found that the transforming gene has no detectable homology to eight retrovirus oncogenes tested.

High molecular weight DNA was prepared from three different types of human sarcoma cell lines and assayed for transforming activity on NIH/3T3 cells using the calcium phosphate co-precipitation technique⁷. Table 1 shows that DNA

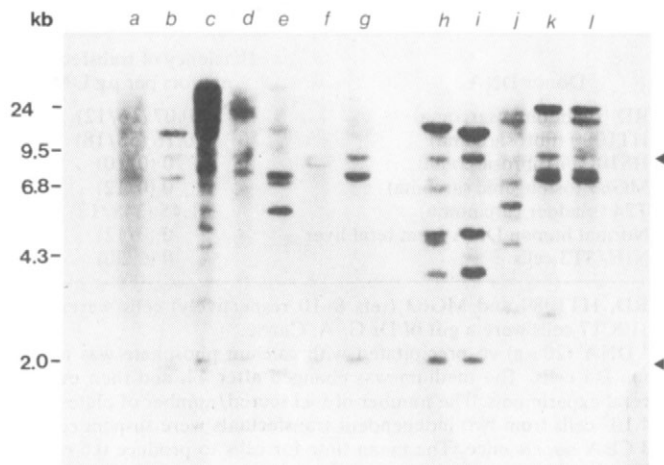


Fig. 1 Presence of human repetitive sequences in secondary transformants derived from RD and HT1080. DNAs (20 µg) were digested with *EcoRI*, phenol-extracted and loaded on a 1-cm thick, 0.8% agarose gel. After electrophoresis the DNA was transferred to nitrocellulose and the blot was hybridized essentially as described by Wahl *et al.*²⁰ except that the formamide concentration was 30%, the hybridization temperature 60°C and mouse DNA was used as carrier. The probe used was fetal liver DNA (previously sonicated to an average length of 4 kb) nick-translated to a specific activity of 2×10^8 c.p.m. per µg. The filter was washed in $0.4 \times \text{SSC}/0.1\%$ SDS for 30 min at 60°C, and subjected to autoradiography with an intensifying screen at -70°C for 5 days. Lanes a, NIH/3T3; b, c, secondaries derived from HT1080 primary 125; d, e, secondaries derived from HT1080 primary 144; f, g, secondaries derived from HT1080 primary 120; h, i, secondaries derived from RD primary 145; j, secondary from RD primary 149; k, l, secondaries derived from RD primary 148. Arrows show the position of the 8.8 kb band conserved in all transfectants and the 1.8 kb band present in some transfectants. (In lane e of the blot shown the 8.8 kb band is poorly visible, however in other blots of this DNA the band is present though faint.)

from embryonal rhabdomyosarcoma cell line RD (ref. 8) and from one of two fibrosarcomas, HT1080 (ref. 9), was able to transform NIH/3T3 cells. DNA extracted from a second fibrosarcoma cell line, HS10CL7, and from an osteogenic sarcoma, MG63 (ref. 10), as well as normal human DNA did not cause morphological transformation. The transforming efficiencies of RD and HT1080 DNAs were similar, though lower than that of the T24 bladder carcinoma cell line. Individual cells in sarcoma DNA-induced foci were morphologically similar to those produced by the bladder line, but the foci themselves were smaller. This suggests that the transfectants produced using sarcoma DNA grow more slowly than those derived using T24 DNA. The morphologically altered cells were found to initiate rapidly growing tumours with short latent periods (10 days) when injected into nude mice whereas the clone of untransformed 3T3 cells did not produce tumours over a 2-month observation period (Table 1). The transfectants differed from the NIH/3T3 cell line in being able to grow in low-serum and in anchorage-independent conditions. Furthermore, DNA isolated from primary transfectants was able to transform NIH/3T3 cells, demonstrating that the transforming activity could be passed serially (Table 1).

To determine whether the transforming activity of the human sarcoma cell lines was related to sequences previously identified in other tumour types or to known viral oncogenes the following experiments were performed.

First, DNAs from the sarcoma cell lines, or from primary transfectants, were digested with various restriction endonucleases before transfection of NIH/3T3 cells. Enzymes having sites in the transforming sequence should destroy the transforming activity of the DNA. Digestion of DNA from primary transfectants with six enzymes that cut DNA on average every 3-7 kilobases (kb) completely destroyed the transforming

Table 1 Transfection of NIH/3T3 cells with DNAs from human sarcoma cell lines

Donor DNA	Efficiency of transfection* (no. foci per μg DNA)	Tumorigenicity of transfectants†	Efficiency of secondary transfection*‡ (no. foci per μg DNA)
RD (rhabdomyosarcoma)	0.07 (16/12)	2/2, 2/2 (10 days)	0.14 (11/4), 0.16 (13/4), 0.13 (10/4)
HT1080 (fibrosarcoma)	0.16 (58/18)	3/3, 3/3 (10 days)	0.13 (10/4), 0.28 (22/4), 0.73 (44/3)
HS10C17 (fibrosarcoma)	0 (0/10)	—	—
MG63 (osteogenic sarcoma)	0 (0/12)	—	—
T24 (bladder carcinoma)	1.45 (348/12)	3/3, 3/3 (10–12 days)	6
Normal human DNA from fetal liver	0 (0/12)	—	—
NIH/3T3 cells	0 (0/30)	0/8 (8 weeks)	—

RD, HT1080 and MG63 (refs 8–10 respectively) cells were obtained from the American Type Culture Collection via Flow Laboratories. HS10C17 cells were a gift of Dr G. A. Currie.

* DNA (20 μg) co-precipitated with calcium phosphate was applied to a 60-mm culture dish seeded 24 h previously with 3×10^5 NIH/3T3 clone D4 cells. The medium was changed after 4 h and then every other day and foci were scored after 16–17 days. Values are the mean of several experiments. The number of foci scored/number of plates examined is given in parentheses.

† 10^6 cells from two independent transfectants were suspended in 0.2 ml of medium and injected subcutaneously into two or three 8–12 week old CBA *nu/nu* mice. The mean time for cells to produce 0.5 cm diameter tumours was also recorded. Each value represents the takes from a different independent primary transfectant.

‡ Each set of values represents the transfection efficiencies for a different primary transfectant.

activity (Table 2). Digestion of DNA from a primary transfectant induced by the bladder carcinoma line T24 showed a different sensitivity to digestion, so the sarcoma transforming sequence is not identical to that of the bladder carcinoma¹¹. Similarly, comparison with published data shows that the transforming sequence of the sarcoma DNAs is not the same as that present in mammary³ or lymphocyte⁶ neoplasms. In addition, none of four enzymes (*Xho*I, *Cla*I, *Sal*I and *Sma*I) having restriction sites that contain the rare dinucleotide CpG and which therefore cut human DNA relatively occasionally inactivated the sarcoma transforming activity. The restriction enzyme sensitivity profiles of RD and HT1080 DNAs are identical for 10 enzymes, suggesting that the transforming activity is associated with the same sequence in both tumours (Table 2).

To characterize the transforming sequence further, we analysed the DNA of secondary transfectants for the presence of human repetitive sequences. After two consecutive rounds of transfection only a small amount of human DNA remains in the transformed cells; any repetitive DNA sequences that are conserved in all secondary transfectants would therefore be expected to be closely linked to or contained in the transforming

Table 2 Effect of restriction endonuclease digestion on the transforming activity of sarcoma DNAs

Enzyme	No. of foci per μg DNA		
	HT1080-f125	RD-f145	T24-f106
No enzyme	0.2	0.28	1.0
<i>Eco</i> RI	0	0	1.0
<i>Bam</i> HI	0	0	ND
<i>Hind</i> III	0	0	0.51
<i>Bgl</i> II	0	0	0.4
<i>Pvu</i> II	0	0	0
<i>Pst</i> I	0	0	0
<i>Xho</i> I	0.1	0.2	ND
<i>Sal</i> I	0.22	0.18	ND
<i>Cla</i> I	0.32	0.18	ND
<i>Sma</i> I	0.13	0.07	ND

DNA (100–200 $\mu\text{g ml}^{-1}$) was digested according to the manufacturers' instructions but using a 10-fold excess of enzyme. After addition of enzyme an aliquot containing 1 μg of DNA was removed and added to 1 μg of λ DNA. Complete digestion of λ DNA was taken to correspond to complete digestion of cellular DNA. 60 μg of digested DNA were added to 120 μg of high molecular weight mouse DNA and calcium phosphate co-precipitates were divided between six plates of NIH/3T3 cells. For all cell lines, DNAs of primary transfectants were used for restriction digests because their transforming activity was higher than that of the original cell lines. In confirmatory experiments, digests of DNA from the original human cell lines gave similar results.

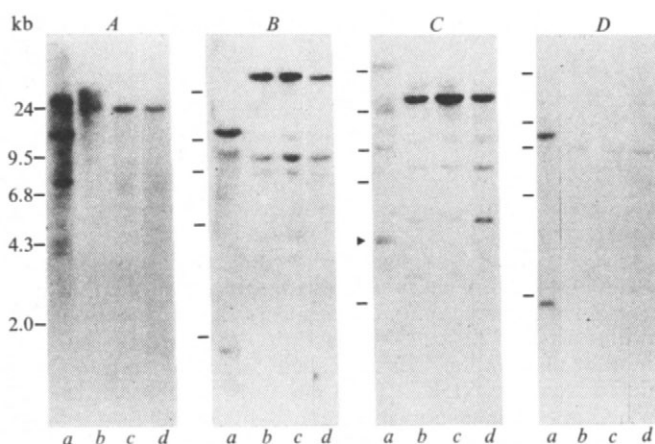


Fig. 2 Southern blots of *Eco*RI (A–C) or *Bam*HI (D) digestions of the following DNAs: a, human fetal liver; b, NIH/3T3; c, a primary HT1080 transfectant; and d, a primary RD transfectant. The blots were hybridized with the following ³²P-labelled viral oncogene probes: A, a 1.5 kb *Pst*I fragment containing *v-myc*; B, a 2.5 kb *Pvu*II fragment containing *v-erb-A* and *v-erb-B*; C, a 1.8 kb *Pvu*II fragment containing *v-Ki-ras* and flanking sequences (the arrow indicates the position of the human 3 kb *Eco*RI fragment homologous to the viral oncogene and identified in NIH/3T3 cells transfected with human lung carcinoma DNA¹⁶; D, a 0.9 kb *Pst*I/*Xba*I fragment containing *v-sis*. All four blots were hybridized in 50% formamide, 5 × SSC at 42 °C except D which was at 50 °C. The filters were washed as follows: A, 0.3 × SSC, 65 °C; B, 0.1 × SSC, 45 °C; C, 0.3 × SSC, 60 °C; and D, 0.6 × SSC, 50 °C.

sequence^{4,5}. A Southern blot¹² profile of such conserved repetitive sequences can be used to distinguish different transforming sequences^{4,5}. Accordingly, DNAs were isolated from secondary transfectants derived from RD and HT1080 DNAs, digested with *Eco*RI and probed on Southern blots with nick-translated total human DNA¹³. In the conditions of hybridization used, only human repetitive sequences can be detected. Although several fragments are present in each transfectant, only a single 8.8 kilobase pair (kbp) *Eco*RI fragment is conserved in all secondary transfectants (Fig. 1). Identical Southern blots probed with the plasmid BLUR-8 (ref. 14), which contains a cloned repetitive sequence from the human *Alu* family, show a similar pattern of hybridization (data not shown). However, the band corresponding to the 8.8 kbp *Eco*RI fragment is less intense than the other bands. We assume that the repetitive sequence present in this fragment is only distantly related to the *Alu* repetitive element contained in BLUR-8. Many of the secondary transfectants also contain a second fragment of 1.8 kbp (Fig. 1). However, as this fragment is not conserved in all secondary transfectants, it is probably not contained in the transforming sequence but may be linked to it. The observation that an 8.8 kbp *Eco*RI fragment is conserved in all secondary transfectants derived from both RD and HT1080 DNAs again strongly suggests that the transforming sequence in the two

DNAs is the same. Furthermore, the size of the *EcoRI* fragment is different from the fragments found by other investigators in secondary transfectants from DNAs of colon, lung, neuroblastoma and promyelocytic leukaemia^{4,5}.

Recently it has been shown that the transforming sequence activated in bladder carcinoma cell lines is a human cellular homologue of *v-Ha-ras*, the oncogene of Harvey sarcoma virus¹⁵. The transforming sequence of a lung carcinoma cell line is homologous to *v-Ki-ras*, the oncogene of Kirsten sarcoma virus¹⁶. Furthermore, Eva *et al.*¹⁷ have reported that human sarcoma cells contain high levels of RNA transcribed from the cellular homologue of *v-sis*, the oncogene of simian sarcoma virus. It was therefore important to determine whether the sarcoma transforming sequence is related to these or other retrovirus oncogenes. We have probed Southern blots of transfectants to see whether they contain human sequences homologous to *v-erb-A*, *v-erb-B*, *v-Ki-ras*, *v-myc*, *v-sis* (Fig. 2) and to *v-abl*, *v-Ha-ras* and *v-src* (data not shown). None of these probes detected any human sequences in the transfectants although the conditions used detected the mouse homologues

in the transfectants and the human homologues in human DNA.

We conclude from these experiments that the same transforming sequence is present in cell lines derived from the two different types of sarcoma, a rhabdomyosarcoma and a fibrosarcoma. (Recently other workers¹⁸ noted that HT1080 DNA has transforming activity.) We have shown that this sarcoma sequence is different from transforming sequences identified in other human tumour lines and in eight RNA tumour viruses. These data therefore support the idea that different transforming sequences are activated in different tumour types and that these sequences may have a role in both the normal physiology and the neoplastic transformation of those tissues^{3-6,19}. Further characterization of the transforming sequence and the elucidation of its mode of action will depend on the isolation of molecular clones of the transforming gene and its normal allele.

We thank James Christie and Micaela Fairman for technical assistance, R. Avery, J. M. Bishop, R. C. Gallo, J. Neil and B. Ozanne for cloned retrovirus oncogenes and C. Schmid for the BLUR-8 clone. Financial support was generously provided by the Cancer Research Campaign and MRC.

Received 21 June; accepted 26 July 1982.

- Shih, C., Padhy, L. C., Murray, M. & Weinberg, R. A. *Nature* **290**, 261-264 (1981).
- Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1181-1184 (1981).
- Lane, M.-A., Saiten, A. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5185-5189 (1981).
- Murray, M. J. *et al. Cell* **25**, 355-361 (1981).
- Perucho, M. *et al. Cell* **27**, 467-476 (1981).
- Lane, M.-A., Saiten, A. & Cooper, G. M. *Cell* **28**, 873-880 (1982).
- Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456-467 (1973).
- McAllister, R. M., Melnyk, J., Finklestein, J. Z., Adams, E. C. & Gardner, M. B. *Cancer* **24**, 520-526 (1969).
- Rasheed, S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P. & Gardner, M. B. *Cancer* **33**, 1027-1033 (1974).
- Billiau, A. *et al. Antimicrob. Ag. Chemother.* **12**, 11-15 (1977).
- Goldfarb, M., Shimizu, K., Peruchio, M. & Wigler, M. *Nature* **296**, 404-409 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Gusella, J. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2829-2833 (1980).
- Rubin, C. M., Houck, C. M., Deininger, P. L. & Schmid, C. W. *Nature* **284**, 372-374 (1980).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474-478 (1982).
- Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
- Eva, A. *et al. Nature* **295**, 116-119 (1982).
- Pulciani, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2845-2849 (1982).
- Weiss, R. A., Teich, N., Varmus, H. E. & Coffin, J. *RNA Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1982).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).

The production of membrane or secretory forms of immunoglobulins is regulated by C-gene-specific signals

Luciana Forni & Antonio Coutinho*

Basel Institute for Immunology, 487 Grenzacherstrasse, Postfach, 4005 Basel, Switzerland

* Department of Immunology, Umeå University, 90185 Umeå, Sweden

Selective patterns of antibody isotypes are produced in response to thymus-dependent and thymus-independent antigens and mitogens¹⁻⁵. Together with information on the organization of immunoglobulin C_H genes in myelomas^{6,7}, various models on the control of C-gene expression and switch in normal B cells have been proposed but only secreted products or secretory cells have been considered. We report here a dissociation between expression of membrane-bound immunoglobulins and secretion of the same isotype, and in this case describe C-gene specific signals which regulate the production of membrane versus secretory forms of immunoglobulin. These results indicate that regulation of isotypic patterns operate at levels other than immunoglobulin gene structure, and suggest that the secretory phenotype alone is inadequate as the measure of C-gene expression.

Spleen cells were exposed to either the T-independent mitogen lipopolysaccharide (LPS)⁸ or helper T cells which specifically activate a large set of B cells⁹. Activation of B cells under these two conditions results in two distinct patterns of γ -isotype secretion although activation of μ secretion is comparable in both cases. T-independent induction results in predominant secretion of $\gamma 2b$ and $\gamma 3$, while helper cells induce primarily $\gamma 1$ and $\gamma 2a$, but here we have also studied the expression of membrane-bound immunoglobulins. As shown in Table 1, following T-independent induction, $\gamma 1$ is the predominant membrane isotype, indicating a clear discrepancy between the expression and the secretion of this particular

isotype. The dissociation between expression of membrane and secretory form of $\gamma 1$ cannot be explained by different kinetics, as high frequencies of $\gamma 1$ -bearing blasts are observed from days 3 to 7 of culture. Passive adsorption of $\gamma 1$ on LPS blasts is unlikely because the concentration of this isotype in culture supernatants remains low (A.C. and L.F., in preparation), and therefore would require the presence of Fc receptors with exquisite specificity and high affinity for $\gamma 1$, which is not the case for B lymphocytes¹⁰. In contrast, no dissociation of membrane and secretory isotypes was observed after exposure to T helper cells.

We addressed the question of the fate of such LPS-induced, $\gamma 1$ -bearing cells, and of the mechanisms responsible for the dissociation between expression and secretion observed in LPS cultures. We attempted to correct the secretory defect of LPS-activated populations by exposure to specific T-cell help. LPS-activated blasts re-exposed to either LPS or to helper cells produce 3 days later quite distinct isotypes (Table 2). The T-dependent re-stimulation resulted in: the appearance of a high number of $\gamma 1$ secretors which are missing in LPS culture (the number of these is very close to the one of $\gamma 1$ -bearing cells in the starting population); decrease of $\gamma 3$ and $\gamma 2b$ secretors (although at values still higher than in helper cultures) and a significant increase of α plaque-forming cells (PFC). Note that in both cases a comparably high fraction (40%) of all recovered cells are high rate secretors and, of these, a similar proportion (57% and 56%) secrete γ ; both stimuli seem therefore to be equally competent in inducing switch and terminal maturation. The differences are isotype specific and relate to the quality of the stimulus controlling terminal maturation in the same population of activated cells.

To define further the specificity of the maturation signals, we analysed similar cultures both for cells expressing membrane-bound immunoglobulins and PFC. Table 3a shows that LPS is fully competent to induce $\gamma 3$ secretion, and ineffective with regards to $\gamma 1$ secretion. On the other hand, helper activity does induce $\gamma 1$ secretion. As, however, it also doubles the number of $\gamma 1$ -bearing blasts, LPS-activated, $\gamma 1$ -bearing blasts may remain nonsecretory in the presence of helper cells, which

Table 1 Isotype distribution of secreted and membrane-bound immunoglobulin molecules

C _H isotype	Isotype distribution (%) [*]			
	Cultures stimulated with LPS		Cultures stimulated with specific helper T cells	
	Secreted [†]	Membrane bound [‡]	Secreted [†]	Membrane bound [‡]
γ3	35.9	21.3	2.6	0.4
γ1	3.3	35.4	66.0	67.2
γ2b	51.1	27.0	4.6	9.1
γ2a	8.1	14.7	19.0	16.2
α	1.3	1.7	7.4	7.1

Normal spleen cells from C3H/Tif mice were cultured in RPMI medium supplemented with 5×10^{-5} M 2-mercaptoethanol, 10 mM glutamine, 10 mM HEPES, antibiotics and 20% fetal calf serum (Gibco) (5×10^5 ml⁻¹, 2 ml⁻¹ per culture). The cultures were stimulated with either 50 µg ml⁻¹ lipopolysaccharide from *Salmonella abortus equi* (Difco) or 2×10^5 helper cells from C3H/HeJ anti-C3H/Tif cell lines that were established and maintained in continuous growth *in vitro* over the last 16 months by weekly stimulation with irradiated C3H/Tif spleen cells, according to methods previously described in detail¹⁹. The helper cells used in these experiments were from clones isolated by limiting dilution in suspension cultures which activate C3H/Tif spleen cells as described, but are MHC-restricted (unpublished), in contrast to uncloned lines reported before in this strain combination^{19,20}. The 'effector helper to target spleen cell' ratios used here (1:5) had previously been found optimal in these conditions.

^{*} The values are expressed as percentages of the total numbers of cells expressing the isotypes listed in the table. These are mean values of two independent experiments.

[†] Assays for secreting cells were performed after 6–7 days of culture when non-IgM PFC accounted for 50–60% of all PFC, representing 20–25% of all viable cells. Plaque-forming cells (PFC) were detected in a Protein A plaque assay using isotype-specific rabbit antisera^{21,22} raised by injection of purified myeloma proteins of the different isotypes.

[‡] Assays for membrane-bound isotypes were performed at days 3–4 of culture. Cells bearing membrane isotypes other than µ accounted for 25–30% of all cells. When tested in double staining with anti-µ, most cells were found also to bear IgM in conjunction with another C_H isotype. Rabbit antisera raised against purified myeloma proteins were made specific by adsorption of Sepharose-coupled myeloma proteins of different isotypes, as well as F(ab')₂ fragment prepared by pepsin digestion of total IgG from pooled mouse sera. The specificity of the antisera was assayed in solid-phase enzyme immunoassay²³. Conjugation with fluorochromes as well as staining for membrane-bound immunoglobulins was performed as described²⁴.

independently induce γ3-bearing cells to switch to γ1 expression and secretion, thus accounting for the disappearance of γ3 producers on the whole. To distinguish between the two possibilities we examined single cells for the presence of membrane-bound and intracytoplasmic immunoglobulins by double immunofluorescence. This was done soon after restimulation because with the half lives of message and proteins, it should be possible to determine the isotype persisting on the membrane of cells that were newly activated to secretion. Table 3b shows that about half of the γ1-bearing blasts, which in this short period had not greatly increased in number, had been activated to produce intracytoplasmic γ1 by helper cells. This resulted in a proportion of secretors, out of the total γ1 population, which is one order of magnitude higher than in parallel LPS

cultures. The switch from γ3 to γ1, measured as cells with intracytoplasmic γ1 and membrane γ3, was proportionally more frequent in LPS than in helper cell-stimulated cultures. On the other hand, the decay of γ3 cells, both membrane-positive and secretors, was observed after 20 h of restimulation. As the proportion of secretors out of the total γ3 population remains unaffected, the decay of γ3-producing cells can only be explained by inability of helper cell signals, in contrast to LPS, to maintain proliferation and/or viability of γ3-bearing cells and to activate them to secretion. Preliminary experiments indicate that comparable results can be obtained with supernatants of the T-cell clones used in these experiments, which suggests that neither of the T-cell effects, induction of γ1 secretion and inability to support γ3 production, requires cell-to-cell contact.

To confirm that γ1 secretors in helper cell-restimulated cultures derive from γ1-bearing blasts present in LPS cultures, positive selection in a fluorescence-activated cell sorter (FACS II) was carried out. As shown in Table 4, a large number of γ1 high-rate secretors were induced within 15 h by T helper cells in the γ1-enriched but not in the γ1-depleted population, indicating that few, if any γ1 PFC are newly derived from either µ or γ3 to γ1 switches. As the majority of the γ1-bearing blasts in LPS cultures switched from µ (and δ) production after mitogen stimulation (only 2% or less of spleen cells are γ1 positive) we conclude that thymus-independent B-cell activation can induce switches to γ1 expression but is very inefficient in activating γ1 secretion. This operates in a strict C-gene specific manner, because the mitogen is fully competent to induce secretion not only of γ3, but also of γ2b and γ2a, which are coded by genes located distally from γ1. Table 4 also shows that helper cells are much less efficient than LPS in inducing γ3 secretion in the γ1-depleted population and this, in view of the short subculture period, is unlikely to be the result of preferential survival or expansion of γ3 cells in LPS cultures.

The present observations show that measurements of antibodies and secreting cells is insufficient when considering C-gene expression. They also point out that isotype commitment in antibody responses^{1–3} is regulated at levels other than DNA rearrangements and gene structure. In the example described here, the regulation of membrane-versus-secretory forms of immunoglobulins is intimately connected to the control of terminal maturation to high-rate secretion and dependent on the quality of the activating signal. It has been known for some time that the carboxy terminus of membrane µ is encoded in a separate exon, distinct from the Cµ4 domain¹¹, and it is now established that all other C_H genes exhibit the same type of structure^{12–14}. Models based on differential nuclear RNA splicing and/or selective termination of transcription, as well as post-translational controls, have been proposed to explain the production of membrane-bound or secretory heavy chains. The physiological mechanisms regulating these processes are unknown, but the present experiments prove the existence of

Table 2 Differential effects of LPS and helper T cells on secretion of γ isotypes

	PFC per 10 ⁵ B cells [*]		γ1	γ2b	γ2a	α	Total PFC	Total non-IgM PFC
	µ	γ3						
Input cells	12,214	292	62	699	211	84	13,462	1,348
(day 3 LPS blasts)	(ND)	(5,900) [†]	(10,400)	(5,800)	(3,200)	(ND)		
Subculture with LPS	15,927	5,236	211	11,345	4,282	44	37,045	21,118
Subculture with TH	18,391	1,242	10,881	7,645	3,434	479	42,072	23,681
Isotype distribution (%)								
Input cells	90	2	0.5	5	2	0.5		
Subculture with LPS	43	14	0.5	31	12	0.1		
Subculture with TH	44	3	26	18	8	1		

C3H/Tif spleen cell cultures were stimulated with LPS and assayed after 3 days of culture for the presence of membrane-bound immunoglobulins by immunofluorescence and of plaque-forming cells of different C_H isotypes (see Table 1 legend). Such 3-day LPS blasts were then subcultured in Microtitre II plates (Falcon) at 2×10^4 cells in 0.2-ml cultures in the same medium with either LPS (50 µg ml⁻¹) or with C3H/HeJ anti-C3H/Tif helper cells (4×10^3 per culture) and isotype-specific PFC were determined after an additional 3 days of culture.

^{*} The values are expressed per 10⁵ initial cells, and per 10⁵ recovered cells, to facilitate the comparison between the results obtained with the different assays. The total number of B cells recovered per culture was very similar after exposure to LPS or to T helper cells (1.8 ± 0.2 and $1.65 \pm 0.18 \times 10^5$ respectively).

[†] Figures in parentheses represent the number of cells in the starting population positive on the membrane by immunofluorescence. ND, not determined.

Table 3 Effect of LPS and T helpers on expression and secretion of $\gamma 3$ and $\gamma 1$

a	LPS		% Blasts in culture		3 days LPS	
	day 3		3 days LPS	3 days LPS	3 days LPS	3 days TH
$\gamma 3$ membrane	5.8		3.1		0.35	
$\gamma 3$ PFC	0.5		5.3		0.45	
$\gamma 1$ membrane	10.4		9.1		20.4	
$\gamma 1$ PFC	0.4		0.35		8.9	

b	LPS		% Blasts in culture		TH	
	day 3	day 4	+1 day LPS	+1 day TH	day 3	day 4
$\gamma 3$ membrane	2.5	6.6	7.0	2.0	0.1	0.1
$\gamma 3$ membrane	ND	3.4	3.6	1.5	ND	0.05
$\gamma 3$ cytoplasm	ND	0.85	0.9	0.95	ND	0.05
$\gamma 1$ cytoplasm	5.7	8.6	7.1	8.1	12.5	15.6
$\gamma 1$ membrane	ND	0.3	0.3	4.3	ND	11.4
$\gamma 1$ cytoplasm						

Proportion of secretors out of the total membrane-positive population						
$\gamma 3$	—	0.51	0.51	0.75	—	—
$\gamma 1$	—	0.035	0.042	0.53	—	0.73

Day 3 LPS blasts subcultured with either LPS or T helper cells (TH) and assayed in the conditions described in the legends to Tables 1 and 2. a, 3 days subculture. Separate samples of the same cultures were stained by immunofluorescence for membrane-bound immunoglobulins, or assayed for $\gamma 3$ and $\gamma 1$ PFC. b, 20 h subculture. The same cell samples were stained by double immunofluorescence for membrane-bound and intracytoplasmic $\gamma 3$ and $\gamma 1$ (ref. 24). ND, not determined.

Table 4 Early effects (15 h subculture) of LPS and T helper cells on secretion of $\gamma 3$ and $\gamma 1$ by selected membrane $\gamma 1^+$ and $\gamma 1^-$ LPS blasts

Cell population	Isotype of PFC	PFC per 10^5 recovered cells		
		Mitogen in subculture		
		None	LPS	T helper cells
Membrane $\gamma 1^+$	$\gamma 1$	915	900	6,450
LPS blasts	$\gamma 3$	375	550	300
Membrane $\gamma 1^-$	$\gamma 1$	<15	<15	300
LPS blasts	$\gamma 3$	650	9,150	2,750

Day 3 LPS-activated blasts were stained with fluorescein-labelled anti- $\gamma 1$ and separated in the fluorescence-activated cell sorter (Becton and Dickinson FACS II). The $\gamma 1^+$ and $\gamma 1^-$ populations represent the 10% brightest and 10% dimmest cells, respectively. Conditions of primary and secondary cultures and assays were as described in the legends to Tables 1 and 2.

such controls and demonstrate that they are C-gene specific.

Our results stress that such C-gene specificity concerns terminal maturation (high rate production of secretory form), and not switch, as implied before¹⁻⁵ and might operate independently of whether the cell has undergone¹⁵⁻¹⁷ a switch recombination or not¹⁸. Whatever the precise mechanism involved here, it is clear and quite astonishing that the ability to produce the secretory form of immunoglobulins is related to the quality of the activating signal in a C-gene specific manner. As we have studied polyclonal B-cell responses that proceed independently of immunoglobulin (V-region) specificity, it is likely that such C-gene specific regulation is mediated by cellular structures other than immunoglobulin receptors.

The Basel Institute for Immunology was founded and is supported by Hoffmann-La Roche, Basel. This work was partly supported by the Swedish MRC. We thank Ms M. Tuneskog for technical help, Ms W. Breisinger and Ms M. Bühler for typing the manuscript, all participants in the Second Umeå Seminar on Immunobiology for discussion, and Chris Paige for critical reading of the manuscript.

Received 29 March; accepted 22 July 1982.

- Pearlmuter, R. M. *et al.* *J. exp. Med.* **149**, 993-998 (1979).
- Slack, J., Der-Balian, G. P., Nahm, M. & Davie, J. M. *J. exp. Med.* **151**, 853-862 (1980).
- Mongini, P. K. A., Stein, K. E. & Paul, W. E. *J. exp. Med.* **153**, 1-12 (1981).
- Martinez-A. C., Coutinho, A. & Augustin, A. A. *Eur. J. Immun.* **10**, 698-702 (1980).
- Martinez-A. C. & Coutinho, A. *Eur. J. Immun.* (in press).
- Kataoka, T., Kawahara, T., Takahashi, N. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 919-923 (1980).
- Roeder, W., Maki, R., Traunecker, A. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 474-478 (1981).
- Andersson, J., Sjöberg, O. & Möller, G. *Transplant. Rev.* **11**, 131-177 (1972).
- Augustin, A. A. & Coutinho, A. *J. exp. Med.* **151**, 587-598 (1980).
- Dickler, H. P. *Adv. Immun.* **24**, 167-214 (1976).
- Early, P. *et al.* *Cell* **20**, 313-319 (1980).
- Tyler, B. M., Cowmann, A. F., Adams, J. M. & Harris, A. W. *Nature* **293**, 406-408 (1981).
- Maki, R. *et al.* *Cell* **24**, 353-365 (1981).
- Tucker, P. W., Liu, C., Mushinski, J. F. & Blattner, F. R. *Science* **206**, 1353-1360 (1980).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
- Takahashi, N., Kataoka, T. & Honjo, T. *Gene* **11**, 117-217 (1980).
- Liu, C.-P., Tucker, P. W., Mushinski, G. F. & Blattner, F. R. *Science* **209**, 1348-1353 (1980).
- Alt, F. W., Rosenberg, N., Casanova, R. J., Thomas, E. & Baltimore, D. *Nature* **296**, 325-331 (1982).
- Coutinho, A. & Augustin, A. A. *Eur. J. Immun.* **10**, 535-541 (1980).
- Coutinho, A., Augustin, A. A., Forni, L. & Fischer Lindahl, K. *J. Immun.* **127**, 21-25 (1981).
- Gronowicz, E., Coutinho, A. & Melchers, F. *Eur. J. Immun.* **6**, 588-590 (1976).
- Bernabé, R. R. *et al.* *Immun. Meth.* **2**, 187-197 (1981).
- Engvall, E. & Perlmann, P. *Immunochimistry* **8**, 871-874 (1971).
- Forni, L. *Immun. Meth.* **1**, 151-167 (1979).

An adenovirus glycoprotein binds heavy chains of class I transplantation antigens from man and mouse

Christer Signäs, Michael G. Katze*, Håkan Persson* & Lennart Philipson*

Department of Microbiology, University of Uppsala, The Biomedical Center, Box 581, S-751 23 Uppsala, Sweden

The successful killing of virus-infected cells by cytotoxic T lymphocytes (CTL) is dependent on the recognition of both a viral product and class I antigens of the major histocompatibility complex (MHC) on the infected cell surface¹⁻⁹. Whether these two entities are found independently on the cell surface and therefore recognized by two different CTL receptors, or whether they are associated together and can therefore be recognized by a single receptor is not known¹⁰. The association between an adenovirus-encoded glycoprotein expressed on the cell surface early after infection and class I antigens has been investigated and it has been found that antisera against class I antigens can co-precipitate the antigen and the viral glycoprotein from an adenovirus-transformed cell line from the Hooded-Lister rat strain^{11,12}. We show here by *in vitro* affinity chromatography and *in vivo* immunoprecipitation that the viral glycoprotein specifically binds to the heavy chain of class I antigens in both man and mouse.

An adenovirus glycoprotein E3/19K (ref. 13), encoded in region 3 of the adenovirus genome, is found soon after infection. When the glycoprotein fraction from cytoplasmic extracts of infected HeLa cells or a human adenovirus-transformed epithelial cell line (293)¹⁴ superinfected with adenovirus type 2 (Ad2) was immunoprecipitated with anti-HLA antisera both the E3/19K and the heavy chain of the HLA were immunoprecipitated in infected but not in uninfected cells (Fig. 1). From counting the radioactivity in the bands it appears that the two moieties are present in almost equimolar amounts and that most of the HLA chains are bound to E3/19K (not shown). However, in productive infection with Ad2 the E3/19K is overproduced and only around 5-10% of the viral glycoprotein

*Present addresses: Molecular Biology and Genetics, Unit of the Graduate School, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA (M.G.K.); Laboratory of Molecular Genetics, National Institute of Child Health, and Human Development, Bethesda, Maryland 20205, USA (H.P.); European Molecular Biology Laboratory, Postfach 10-2209, D-6900 Heidelberg, FRG (L.P.).

is associated with the class I antigens but in transformed rat cells more than 80% of the viral E3/19K is associated with the class I antigen¹¹. Control experiments showed that artificial mixtures of purified E3/19K and cells did not result in co-precipitation of the complex. The E3/19K protein can thus bind strongly to human transplantation antigen *in vivo* as previously demonstrated for the class I antigen of rats^{11,12}. Given these results we analysed whether this complex could also be formed *in vitro*. Purified E3/19K (Fig. 2 h) was coupled to CNBr Sepharose 4B. The glycoprotein fraction from ³⁵S-methionine-labelled cytoplasmic extracts from two human established lymphoid cell lines (Ingemar B7 and IM-199), which express mainly the class I antigens B7 and A2, were passed over the column. A column prepared with coupled bovine serum albumin (BSA) was the control. Figure 2 shows that a polypeptide of the same molecular weight as the heavy chain of the HLA-B7 allotype specifically bound to the E3/19K column but not to the BSA column. This polypeptide represents the heavy chain of the HLA as revealed by immunoprecipitation (Fig. 2e). It can also be seen that the amount of this polypeptide

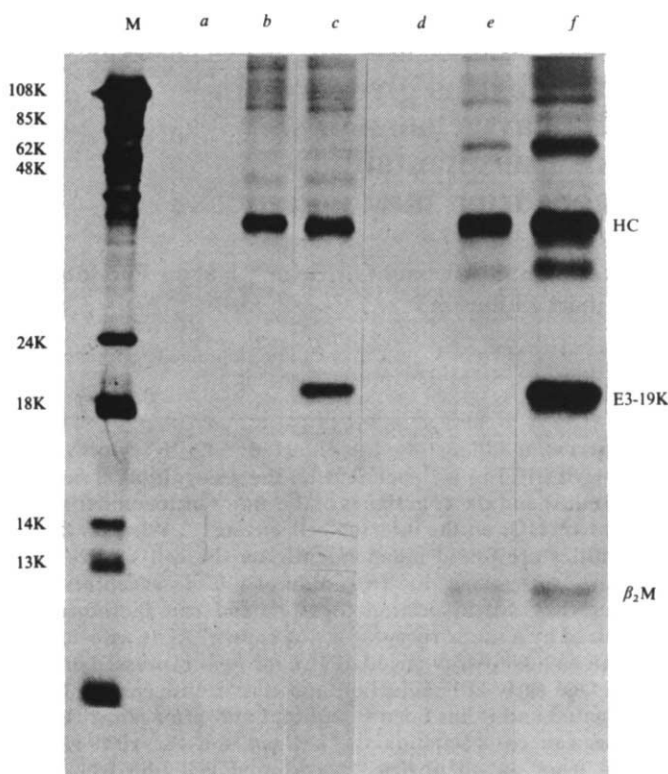


Fig. 1 Immunoprecipitation from uninfected and adenovirus-infected HeLa and 293 cells with anti-HLA serum. Suspension cultures of HeLa cells were labelled overnight with ³⁵S-methionine (10 μ Ci ml⁻¹), to enhance labelling of the class I MHC antigens, in methionine-free spinner medium (M) supplemented with 2.5% normal medium and 7% calf serum. Monolayers of 293 cells were labelled similarly using Eagle's medium and 10% fetal calf serum. The cells were washed twice and infected with adenovirus type 2 (ref. 18). Cycloheximide (25 μ g ml⁻¹) was added 2 h post-infection and removed 3 h later in order to enhance early viral mRNA accumulation¹⁹. The cells were then labelled for an additional 3 h with ³⁵S-methionine (10 μ Ci ml⁻¹) in the presence of cytosine arabinoside (25 μ g ml⁻¹). The cells were washed in phosphate-buffered saline (PBS) and lysed in 0.5% NP40, 150 mM KCl, 1.5 mM MgCl₂, 10 mM Tris-HCl pH 8.0. The cytoplasmic extracts (2 ml from 2 $\times$ 10⁶ cells) were passed over a lentil lectin Sepharose column (4 ml) (Pharmacia) equilibrated with PBS, 0.1% Triton X-100, 1 mM phenylmethylsulfonyl fluoride (PMSF) (buffer A). Elution was performed with buffer A containing 0.1 M α -methylmannoside. The glycoprotein eluate was immunoprecipitated after preprecipitation with an unrelated immune complex and analysed on 13% SDS-polyacrylamide gels²⁰. The anti-HLA serum used was a mixture of a monoclonal antiserum that recognizes HLA-A, -B or -C antigens²¹ and a rabbit antiserum made against papain-solubilized HLA. Lanes a, d, immunoprecipitation with normal rabbit serum; b, c, e and f, immunoprecipitation with anti-HLA serum. a, c, infected HeLa cells; b, uninfected HeLa cells. d, f, infected 293 cells; e, uninfected 293-cells. In this and subsequent figures ³⁵S-labelled adenovirus proteins were used as size markers M.

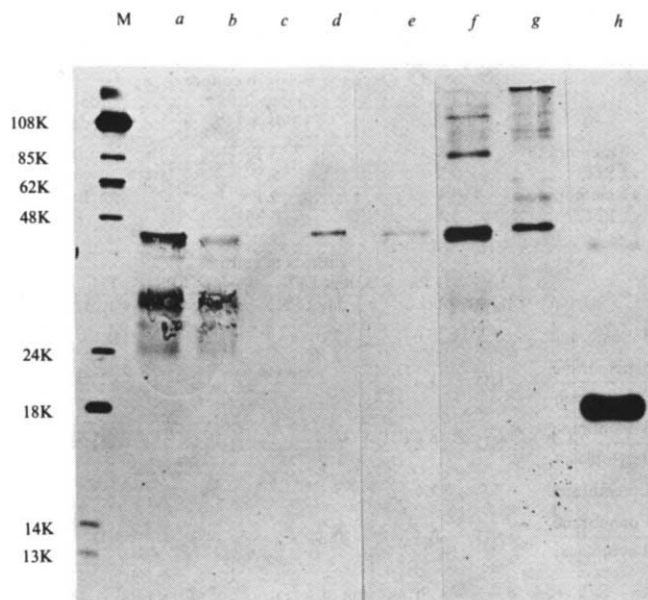


Fig. 2 Affinity chromatography of human class I antigens on the adenovirus glycoprotein. The adenovirus E3/19K protein (100 μ g) or bovine serum albumin (100 μ g) was coupled to 1 ml cyanogen bromide-activated Sepharose 4B (Pharmacia). Two different lymphoid cell lines, HLA-A2 and HLA-B7, were labelled overnight with ³⁵S-methionine (10 μ Ci ml⁻¹), cytoplasmic extracts were prepared, and the glycoproteins isolated by lentil lectin chromatography as described in Fig. 1 legend except that 2 mM dithiothreitol (DTT) was included in the buffers. The lentil lectin eluate was dialysed against buffer A (see Fig. 1), divided in two aliquots and passed over the BSA and the E3/19K columns respectively at 4 °C. The columns were washed thoroughly and eluted with buffer A containing 2 M NaCl and 2 mM DTT. Each fraction was analysed by SDS-polyacrylamide gel electrophoresis and immunoprecipitated as described in Fig. 1 legend. a, Percolate of HLA-B7 from the BSA column; b, percolate of HLA-B7 from the E3/19K column; c, eluate BSA column; d, eluate E3/19K column; e, eluate from E3/19K column immunoprecipitated with anti-HLA serum. No precipitate was observed with normal serum. f, g Show the percolate and eluate fractions of HLA-A2 glycoproteins; h, purified E3/19K. The size marker M is described in Fig. 1 legend.

is decreased in the eluate of the E3/19K column compared with the BSA column. The same results were obtained with the HLA-A2 cells (Fig. 2f, g). These results suggest that the E3/19K can specifically bind to different allotypes of the class I transplantation antigens of man.

We also examined *in vitro* binding of mouse class I antigens using the cell line YAC1 (H-2^a derived from a H-2^k/H-2^d hybrid). Figure 3 shows that a polypeptide of a similar size to that from humans specifically bound to the E3/19K column. Immunoprecipitation with anti-H-2-serum established that this polypeptide corresponded to the heavy chain of the H-2 antigen (not shown). It appears therefore, that the binding of E3/19K to the heavy chain class I transplantation antigens does not show species specificity.

To localize the site of interaction between the two molecules we attempted to perform enzymatic digestions of plasma membranes in order to release the complex. The heavy chain of the class I antigens can be released from membranes with papain, creating the 32,000 (32K) and 34,000 (34K) molecular weight (M_r) subspecies¹⁵. However, papain treatment of infected HeLa cells could not release a complex containing both moieties when analysed by immunoprecipitation with anti-HLA serum. Instead the viral glycoprotein remains in the plasma membrane pellet together with uncleaved or partially cleaved antigen (Fig. 4). The use of an affinity chromatography column with papain released fragments of the human HLA-B7 antigen confirmed this. Figure 4b demonstrates that very few papain fragments are bound to the E3/19K column even though the columns were loaded with material carrying eight times more radioactivity than in the previous experiments. These results suggest that the major interaction occurs either close to the cell membrane or on the cytoplasmic side of the membrane but a

configurational change in the papain fragments cannot be excluded. It is also possible that nonspecific hydrophobic interactions occur between the proteins where they penetrate the membrane, but we consider this unlikely as Triton X-100 was used in all the buffers and a nonspecific interaction may not give the selectivity observed.

If a cytoplasmic interaction guides the complex formation, considerable variation in this region of the transplantation antigens could account for the diversity in CTL recognition. Recent analyses of the genomic structure of the HLA and H-2 genes indeed suggest that the cytoplasmic portions of the class I transplantation antigens are complex and contain at least three exons with varying splice patterns for the H-2 genes^{16,17}.

The strong complex formation between adenovirus glycoprotein and the transplantation antigens appears to be unique among the viral systems investigated so far. The influenza virus haemagglutinin⁹, the vesicular stomatitis virus glycoprotein⁵ and several other identified virus determinants for the T-cell mediated cytotoxicity⁵⁻⁹ do not appear to be co-precipitated with the heavy chain of the class I antigens. The adenovirus system, including the mouse adenovirus, where an S-antigen is firmly complexed to the class I transplantation antigens of the mouse⁸, may therefore be a model for the altered self hypothesis of virus-induced T-cell cytotoxicity. We have previously demonstrated that this complex in rat cells is recognized by syngeneic rat cytotoxic T cells and that this reaction can be inhibited by syngeneic antibodies to the cells and allogeneic

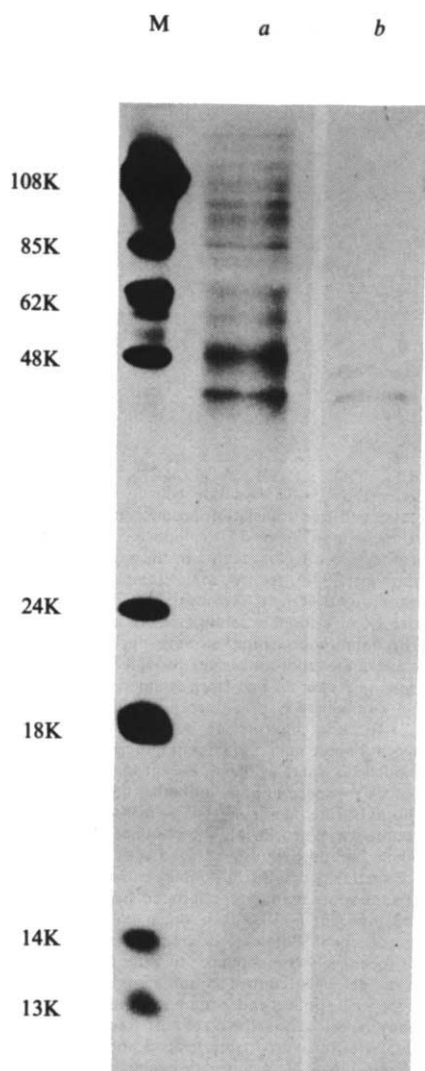


Fig. 3 Affinity chromatography of mouse class I antigens on the adenovirus glycoprotein. A labelled cytoplasmic extract from the YAC (H-2^{kd}) cells was treated exactly as in Fig. 2. *a*, Percolate from the E3/19K column; *b*, eluate. The size marker M is described in Fig. 1 legend.

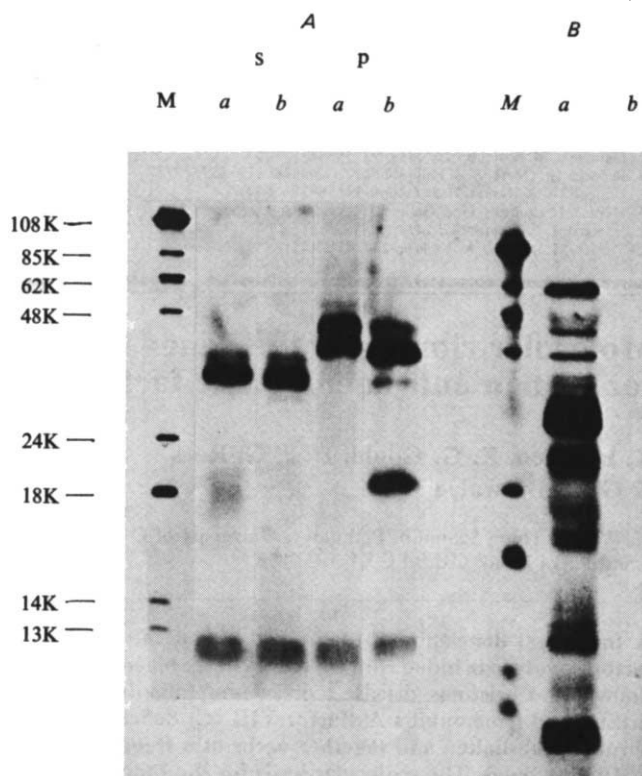


Fig. 4 Immunoprecipitation of papain-treated membranes from infected and mock-infected HeLa cells (*a*) and affinity chromatography of papain-released HLA-B7 on the E3/19K column (*b*). HeLa cells were labelled and infected as described in Fig. 1 legend. The cells were collected by centrifugation and membrane preparation and papain treatment was performed as previously described²². The supernatant after papain digestion was dialysed against PBS, 1% Triton X-100, 1 mM PMSF and immunoprecipitated. The pellet was dissolved in PBS, 1% Triton X-100, 1 mM PMSF by stirring overnight at 4°C. The dissolved pellet was also immunoprecipitated and the samples were analysed by SDS-gel electrophoresis. HLA-B7 cells were labelled as described in Fig. 1 legend. Membrane preparation and papain treatment was performed as described above. The supernatant after papain treatment was dialysed against PBS, 0.1% Triton X-100, 1 mM PMSF and added to the E3/19K column was described in Fig. 2 legend. *A*, Immunoprecipitation with anti-HLA serum from the supernatant (*s*) and pellet (*p*) after papain digestion of HeLa cells; *a*, uninfected; *b*, infected cells. *B*, Affinity chromatography on E3/19K protein column. Lane *a*, percolate, and lane *b*, high salt eluted fraction. The size marker M is described in Fig. 1 legend.

antibodies to the class I transplantation antigen¹². By using the complexes, which are retained on the affinity columns, as immunogen for the production *in vivo* of both specific antibodies and cytotoxic T cells, we hope to be able to define the T-cell receptor and the role of the CTL in virus infection.

We thank B. M. Johansson for technical assistance and Marianne Gustafson for secretarial help. The HLA and H-2 antisera were provided by Drs W. Bodmer and P. Peterson. We also thank Dr J. Andersson, for the YAC-1 cell line and critical comments. M.G.K. was supported by a postdoctoral fellowship from the EMBO. This work was supported by grants from the Swedish MRC and the Swedish Society against Cancer.

Received 5 April; accepted 5 July 1982.

1. Paul, W. E. & Benacerraf, B. *Science* **195**, 1293-1300 (1977).
2. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51-177 (1979).
3. Doherty, P. C., Blanden, R. V. & Zinkernagel, R. M. *Transplant. Rev.* **29**, 89-123 (1976).
4. Langman, R. E. *Rev. Physiol. Biochem. Pharmac.* **81**, 1-37 (1978).
5. Hale, A. H., Witte, O. N., Baltimore, D. & Eisen, H. N. *Proc. nat. Acad. Sci. U.S.A.* **75**, 970-974 (1978).
6. Lawman, M. J. P. *et al. Infect. Immunity* **30**, 451-461 (1980).
7. Finberg, R., Weiner, H. L., Burakoff, S. J. & Fields, B. N. *Infect. Immunity* **31**, 646-649 (1981).
8. Inada, T. & Uetake, H. *Microbiol. Immun.* **24**, 525-535 (1980).
9. Effros, R. B., Doherty, P. C., Gerhard, W. & Bennink, J. J. *exp. Med.* **145**, 557-568 (1977).
10. Zinkernagel, R. M. & Doherty, P. C. *Cold Spring Harb. Symp. quant. Biol.* **41**, 505-510 (1977).
11. Persson, H., Kvist, S., Östberg, L., Peterson, P. A. & Philipson, L. *Cold Spring Harb. Symp. quant. Biol.* **44**, 509-517 (1980).

12. Kvist, S., Östberg, L., Persson, H., Philipson, L. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5674–5678 (1978).
13. Persson, H., Jansson, M. & Philipson, L. *J. molec. Biol.* **136**, 375–394 (1980).
14. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. J. *gen. Virol.* **36**, 59–72 (1977).
15. Krangel, M. S., Orr, H. T. & Strominger, J. L. *Scand. J. Immun.* **11**, 561–571 (1980).
16. Malissen, M., Malissen, B. & Jordan, B. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 893–897 (1982).
17. Steinmetz, M. *et al. Cell* **25**, 683–692 (1981).
18. Persson, H., Öberg, B. & Philipson, L. *J. Virol.* **28**, 119–139 (1978).
19. Craig, E. A. & Raskas, H. J. *J. Virol.* **14**, 26–32 (1974).
20. Hérissey, J., Courtois, G. & Galibert, F. *Nucleic Acids Res.* **8**, 2173–2192 (1980).
21. Barnstable, C. J. *et al. Cell* **14**, 9–20 (1978).
22. Turner, M. J. *et al. J. biol. Chem.* **250**, 4512–4519 (1975).

Molecular cloning of the gene for human anti-haemophilic factor IX

K. H. Choo, K. G. Gould, D. J. G. Rees & G. G. Brownlee

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

A functional deficiency of factor IX, one of the coagulation factors involved in blood clotting¹, leads to the bleeding disorder known as Christmas disease², or haemophilia B. Both this disease and haemophilia A (factor VIII (C) deficiency) are X chromosome-linked and together occur at a frequency of ~1 in 10,000 males. The molecular basis for the functional alteration of factor IX in Christmas disease is not clearly understood. As a first step towards the elucidation of the molecular events involved, we have attempted molecular cloning of the factor IX gene. We used a bovine factor IX cDNA clone, isolated using synthetic oligonucleotides as probes, to screen a cloned human gene library. Here we report the isolation and partial characterization of a λ recombinant phage containing the human factor IX gene.

As a first step in the isolation of a cDNA clone for bovine factor IX, bovine liver was used to prepare total mRNA by the guanidine hydrochloride method³. A preparation enriched in factor IX-specific mRNA was obtained by oligo (dT)-cellulose chromatography, followed by two successive separations by sucrose gradient centrifugation. The enrichment was monitored by a specific immunological assay based on cell-free translation using rabbit reticulocyte lysates (Fig. 1). A sucrose gradient fraction sedimenting at 20–22S was found to be enriched (about 10-fold) for the bovine factor IX mRNA, and was therefore used in the following cloning experiments.

Synthetic oligonucleotides were used⁴ to produce a cloned cDNA library. From the available amino acid sequence data for bovine factor IX⁵, two different sets of oligonucleotides were synthesized. The first set [designated oligo (N-1)] was deduced from the amino acid sequence residues 348–352 (His-Met-Phe-Cys-Ala) and was synthesized, using the solid phase phosphotriester method⁶, as a mixture of eight different sequences (5'-GC⁺CA⁺AAACAT⁺TC-3'). The second set of oligonucleotides [oligo (N-2)] was deduced from amino acids 70–75 (Glu-Cys-Trp-Cys-Gln-Ala); this was synthesized as two mixtures of eight different sequences each (5'-GCTTG⁺CACCA⁺CA⁺TC-3' and 5'-GCCTG⁺CACCA⁺CA⁺TC-3').

From a knowledge of the concentration of factor IX in blood plasma (~2–4 $\mu\text{g ml}^{-1}$) and of the amount of ³⁵S-cysteine counts precipitable by anti-factor IX antibodies, the abundance of factor IX mRNA was estimated to be as low as 0.01%. Therefore, to achieve further enrichment for factor IX mRNA, oligo (N-1) was used as a primer to synthesize cDNA from the sucrose gradient-enriched poly(A)-mRNA preparation. A library of 7,000 colonies was prepared and screened using ³²P-labelled oligo (N-2) (see Fig. 2 legend). One colony (designated BIX-1) gave a positive hybridization signal above background. Characterization of this colony by restriction endonuclease cleavage

(data not shown) indicated that it contained a DNA insert of ~430 base pairs (bp). Figure 2 shows part of the nucleotide sequence, that is, 304 residues of the BIX-1 insert, and confirms that it is a bovine factor IX sequence. It is interesting that only part of the insert sequence (amino acid residues 52–139) corresponds directly to the nucleotide sequence predicted from the known bovine amino acid sequence data⁵. Over this region, there are no discrepancies between BIX-1 and the published data for factor IX, except at nucleotides 38–40 which code for Asp instead of Thr. This amino acid change was also observed in a second, independent cDNA clone. (This latter clone, which

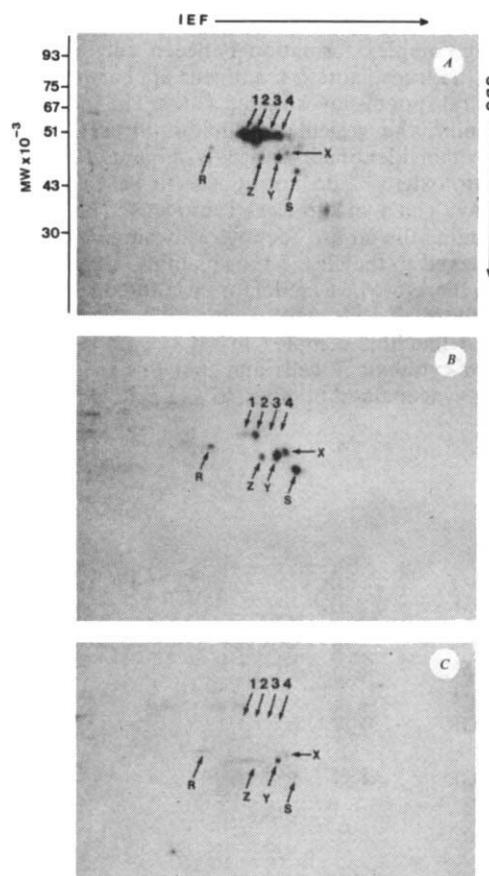


Fig. 1 Two-dimensional polyacrylamide gel electrophoretic analysis of immunoprecipitated cell-free translation products of bovine liver mRNA. mRNA purified through an oligo (dT)-cellulose column¹¹ was translated in a rabbit reticulocyte cell free system¹² in the presence of ³⁵S-cysteine. After incubation at 29°C for 90 min, samples were treated for immunoprecipitation as described previously¹³. Crude antiserum against pure bovine factor IX was raised in rabbit. Specific anti-factor IX immunoglobulins used for immunoprecipitation were purified as described previously¹⁴ by passage of the crude antiserum through a Sepharose-4B column to which pure bovine factor IX had been coupled. After incubation, the antibody-antigen complex was precipitated by the addition of protein A-Sepharose CL-4B resin (Pharmacia), washed and resolved on a two-dimensional polyacrylamide gel¹⁵. A, immunoprecipitation in the presence of purified antibodies; B, same as for A, except that immunoprecipitation was performed in the presence of both antibodies and a competing amount of unlabelled pure bovine factor IX; C, same as for A, except that immunoprecipitation was performed in the presence of a control antiserum from a rabbit which had not been immunized with factor IX. 1–4, Factor IX polypeptide spots; evidence for this comes from the fact that (1) these are the predominant spots immunoprecipitated by the specific antibody; (2) their molecular weight (~50,000, a single polypeptide chain plus a possible pre-peptide signal sequence) is compatible with published data⁵; (3) the spots are specifically competed for by pure factor IX; and (4) they are absent from the gel when control serum is used. (Partial chymotryptic digestion of polypeptide spots 1 and 2 (data not shown) suggests that they are related and may have arisen as a result of post-translational modification; the amount of material recovered from spots 3 and 4 was insufficient for these to be compared.) R, S, X, Y, and Z designate nonspecific background polypeptide spots which form useful reference points on the two-dimensional gel. The molecular weights (MWs) of the polypeptide markers on the second dimension gel are shown on the left. IEF designates isoelectric focusing in the first dimension¹³. SDS, polyacrylamide gel electrophoresis in the second dimension¹³.

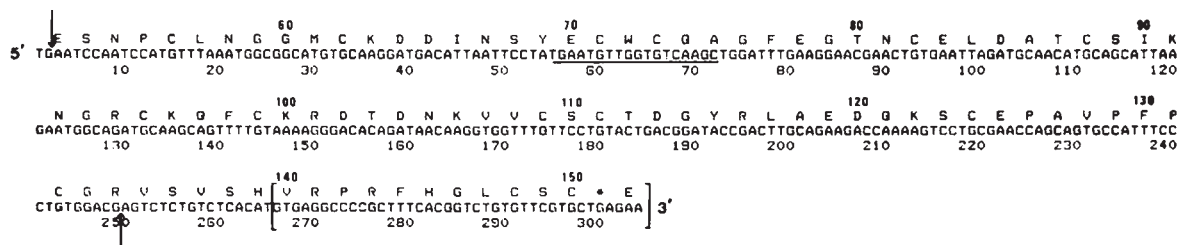


Fig. 2 Isolation and characterization of bovine BIX-1 clone. A library of clones containing bovine cDNA inserts was prepared as follows. First, cDNA strands were synthesized as described elsewhere¹⁵ except that 2 µg of oligo (N-1), 20–30 µg of sucrose gradient-enriched poly (A)⁺ mRNA, 10 µCi [α -³²P]dATP (3,000 Ci mmol⁻¹; Amersham) and 50 U of reverse transcriptase were used in a 50 µl reaction mixture. The cDNA was phenol-extracted, desalted on a Sephadex-G100 column, then treated with alkali (0.1M NaOH, 1mM EDTA) at 60 °C for 30 min to remove the mRNA strand. Second strand DNA was synthesized as described elsewhere¹⁵. The double-stranded DNA was then cleaved with the restriction enzyme *Mbo*I and ligated to pBR322 which had been cut with *Bam*HI and treated with calf intestinal alkaline phosphatase to minimize vector self-religation. The ligated DNA was used to transform *Escherichia coli* strain MC1061 (ref. 16). A library of 7,000 ampicillin-resistant colonies was obtained. Of these, ~85% were tetracycline-sensitive. The library of colonies were transferred in an unordered fashion to 13 Whatman 541 filter papers, amplified with chloramphenicol, and prepared for hybridization as described by Gergen *et al.*¹⁷. The filters were prehybridized at 65 °C for 4 h in 6 × NET, 5 × Denhardt's, 0.5% NP-40 and 1 µg ml⁻¹ yeast RNA described by Wallace *et al.*⁴. Hybridization was at 47 °C for 20 h in the same solution, containing 3 × 10⁵ c.p.m. (0.7 ng) ml⁻¹ of labelled oligo (N-2) probe. Labelling was achieved by phosphorylating the oligonucleotides at the 5' hydroxyl end using [γ -³²P]ATP and T₄ phosphokinase (Boehringer)¹⁵. At the end of the hybridization, filters were washed successively at 0–4 °C (2 h), 25 °C (10 min), 37 °C (10 min) and 47 °C (10 min). On fluorography of the filters from this screening, one colony showed a positive signal above background; this clone was designated BIX-1 and part of the DNA insert was characterized by sequencing performed according to Maxam and Gilbert⁸. Only the coding strand is shown and this is numbered in the opposite orientation to the tetracycline resistance gene of the plasmid. The deduced amino acid sequence (top row) has been numbered according to Katayama *et al.*⁵. The 5' *Sau*3AI cloning site in the BIX-1 clone is eight nucleotides to the left of the start of the above sequence. The arrows indicate two *Hinf*I sites which give rise to a 247-nucleotide fragment (see text). Brackets indicate the start of a sequence which does not correspond to published amino acid data⁵ (see text). The region corresponding to the oligo (N-2) probe is underlined. * Designates a nonsense codon. IXNET is 0.15 M NaCl, 1 mM EDTA, 15 mM Tris-chloride, pH 7.5.

has a DNA insert spanning nucleotides 7–108 as shown in Fig. 2, was isolated using the procedure described for the BIX-1 clone, except that oligo (dT)_{12–18} was used to initiate cDNA synthesis, and the cloned cDNA library was prepared following the homopolymer tailing method of Roychoudhury and Wu⁷. The origin of the adjacent sequence starting at residue 140 (see Fig. 2) which codes for an amino acid sequence unrelated to factor IX is not clear.

A DNA fragment (see Fig. 3 legend) was isolated from the bovine BIX-1 clone and used as a hybridization probe to screen a cloned human gene library; a positive clone, λ HIX1b, was isolated and characterized by restriction endonuclease cleavage and Southern blotting (see Fig. 3). From these results, a partial restriction map for λ HIX1b was constructed (Fig. 4A). A segment of the DNA insert containing the region that hybridizes to the bovine cDNA probe, was identified and isolated for

direct sequencing⁸. The results (Fig. 4B) reveal that, for the 96 nucleotides determined, there is 85% nucleotide sequence homology between λ HIX1b and bovine factor IX. Comparison of the amino acid sequences deduced from the nucleotide sequences shows 78% homology between the two species; note that three of the nucleotide changes are not accompanied by an amino acid change and are therefore silent changes. The degree of homology is compatible with the cross-hybridization observed between the bovine factor IX cDNA probe and the λ HIX1b clone in the rather stringent conditions used, and provides direct evidence that the latter carries an authentic DNA sequence coding for the human factor IX gene.

The present data are equivocal as to whether the λ HIX1b clone carries a complete factor IX gene. The restriction map (Fig. 4A) shows however, that the cloned segment includes more than six kilobases (kb) of DNA on one side and >10 kb

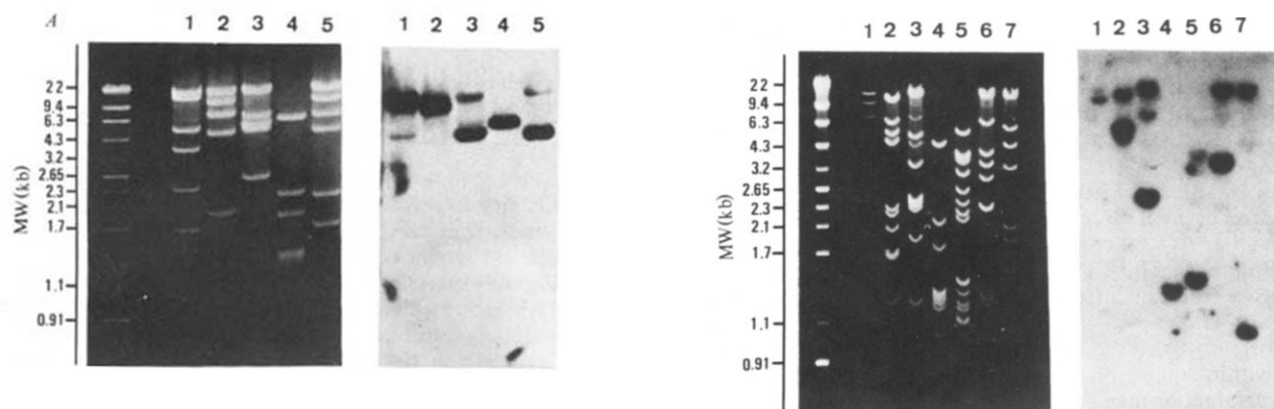


Fig. 3 Isolation and characterization of the human λ HIX1b clone. The cloned human DNA library used was a *Hae*III/*Alu*I λ Charon 4a library prepared by Lawn *et al.*¹⁸. Phage recombinants (10⁶) from this library were screened using the *in situ* plaque hybridization procedure described previously¹⁸. Prehybridization and hybridization were carried out at 42 °C in 50% formamide. After hybridization, filters were washed at room temperature with 2 × SSC, 0.1% SDS, then at 65 °C with 1 × SSC, 0.1% SDS. A 325-nucleotide fragment from the bovine BIX-1 cDNA clone was initially used as a probe in the hybridization. This fragment corresponds to nucleotides –8 to 317 as shown in Fig. 2, and was isolated by *Sau*3AI digestion of BIX-1 plasmid DNA. The isolated DNA was labelled to high specific activity by incorporation of [α -³²P]ATP using Amersham's nick-translation kit; 10 clones were isolated. These were plaque-purified and re-screened with a 247-nucleotide fragment from BIX-1 clone. This fragment, derived from nucleotides 3–249 (see Fig. 2), contains only sequences corresponding to the published bovine factor IX amino acid sequence⁵, and was isolated by *Hinf*I digestion of BIX-1 plasmid DNA. Only a single clone gave a positive hybridization signal with this probe. This clone was further plaque-purified and the resulting clone (designated λ HIX1b) characterized by digestion with various restriction endonucleases. Ethidium bromide-stained gels are shown on the left and their corresponding fluorograms after Southern transfer¹⁹ and hybridization with factor IX-specific 247-nucleotide probe are shown on the right. A, 1–5, cleavage pattern with *Bam*HI, *Bgl*II, *Hind*III, *Hpa*II and *Eco*RI, respectively. The sizes of the fragments showing predominant hybridization signals for these endonucleases are ~16, 11, 5.5, 6.5 and 5.5 kb respectively. B, 1, cleavage with *Bgl*II alone; 2–7, cleavage with *Bgl*II in combination with *Bam*HI, *Hind*III, *Hpa*II, *Pvu*II, *Sst*I and *Eco*RI, respectively. The sizes of the fragments showing predominant hybridization signals for these endonucleases (1–7) are ~11, 5.5, 2.5, 1.4, 1.5, 3.4 and 1.0 kb, respectively. The MW markers (kb) shown on the left are fragments from pBR322 plasmid DNA cleaved with *Alu*I, *Bam*HI + *Pvu*II and *Bam*HI + *Pst*I, plus λ DNA cleaved with *Hind*III.

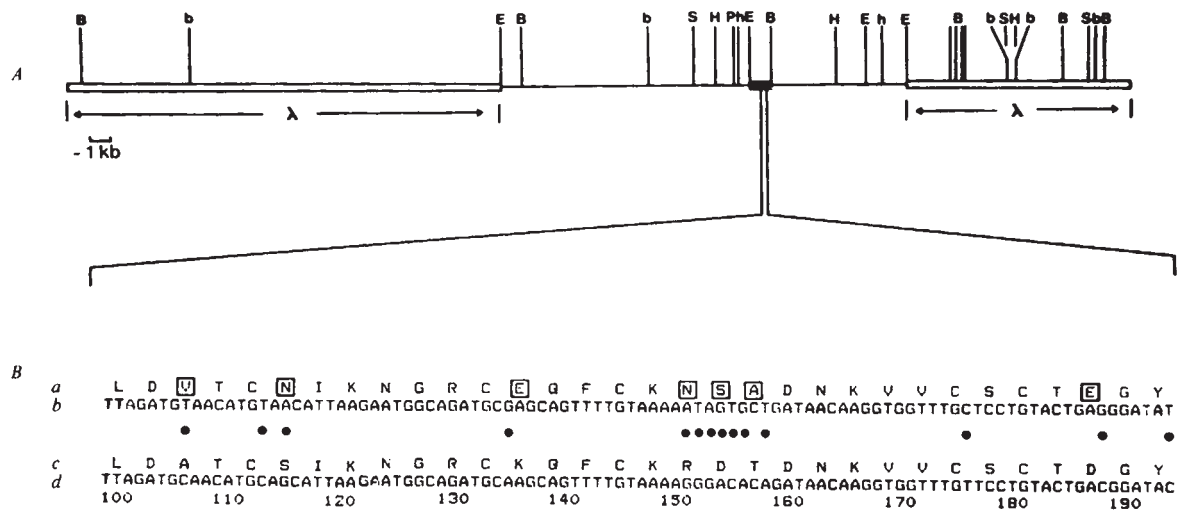


Fig. 4 A, partial restriction map of λHIX1b phage DNA. B = *Bgl*II; b, *Bam*HI; E, *Eco*RI; S, *Sst*I; H, *Hind*III; h, *Hpa*II, P, *Pvu*II. The Charon 4A λ vector long arm is on the left and the short arm is on the right. Restriction sites for *Bgl*II, *Bam*HI, *Eco*II, *Sst*I and *Hind*III on the vector arms are shown. The smallest fragment (~1 kb) within the human DNA insert which hybridizes to the bovine cDNA probe (Fig. 3) is a *Bgl*II-*Eco*RI fragment (indicated by a solid box). The construction of this map was based on the data shown in Fig. 3 and that derived from restriction analysis of a pBR322 subclone of the 10-12-kb *Bgl*II fragment containing the 1-kb *Bgl*II-*Eco*RI probe region. This map shows the approximate positions of the nearest restriction sites to the left and right of the probe region (with the exception of *Sst*I and *Pvu*II where their sites to the right have not been determined). B, Sequence of 96 nucleotides located within a region of the human DNA insert which hybridizes to the bovine cDNA probe. a, Amino acid sequence deduced from the nucleotide sequence b, for λHIX1b. These two sequences are compared with the amino acid c and nucleotide d sequence of bovine factor IX. The bovine nucleotide sequence has been numbered similarly to that shown in Fig. 2. ● Indicates a nucleotide difference between the human and bovine sequences. □ Indicates a difference in amino acid between the two species. Note that the orientation of this sequence within the genomic clone is 5' to 3' from right to left.

on the other side of the sequenced region (or the 'connecting peptide region')⁵, and suggests that this segment potentially contains a complete human factor IX gene. Further sequencing should clarify this and also elucidate the structure of this gene.

Recent reports^{9,10} have described the use of specific gene probes for prenatal and antenatal diagnosis of genetic disorders. These methods are based on the observation of the existence of polymorphisms for restriction endonuclease sites which are closely linked to a particular gene locus. It is hoped that the genomic clone reported here will be similarly useful for the diagnosis and understanding of the mutations in patients with haemophilia B, as well as of heterozygote mothers who are carriers of the factor IX mutation.

We thank Dr M. P. Esnouf for gifts of the crude anti-bovine factor IX antiserum and pure bovine factor IX; Drs F. E. Baralle, N. J. Proudfoot and F. Giannelli for their help and useful discussion; and Dr M. C. Carrol for providing the *Bam*HI-cut and phosphatase-treated pBR322 vector. This work was supported in part by MRC project grant G8103835 to G.G.B. K.H.C. is a recipient of Uncle Bobs Travelling Scholarship and a C. J. Martin Research Fellowship from the National Health and MRC of Australia.

Note added in proof: Further sequence of clone λHIX1b suggests that there is a splice point after residue 101 of the sequence (Fig. 4B, line d). The TTAG (nucleotides 98-101) may therefore derive from an intron and may not code for the amino acids L and D.

The origin of the sequence unrelated to factor IX mRNA shown within square brackets in Fig. 2 could reflect either a cloning artefact or may derive from an intron region assuming we have cloned a region of a pre-mRNA molecule. Further sequencing of the genuine clone should resolve this.

Received 12 May; accepted 23 June 1982.

- Jackson, C. M. & Nemerson, Y. A. *Rev. Biochem.* **49**, 765-811 (1980).
- Ratnoff, O. D. in *The Metabolic Basis of Inherited Disease* (eds Stanbury, J. B., Wyngaarden, J. B. & Frederickson, D. S.) 1755-1791 (McGraw-Hill, New York, 1978).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochem.* **18**, 5294-5299 (1979).
- Wallace, R. B. *et al. Nucleic Acids Res.* **9**, 879-894 (1981).
- Katayama, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 4990-4994 (1979).
- Duckworth, M. L. *et al. Nucleic Acids Res.* **9**, 1691-1706 (1981).
- Roychoudhury, R. & Wu, R. *Meth. Enzym.* **65**, 43-62 (1980).
- Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1980).
- Kan, Y. W., Lee, K. Y., Furbetta, M., Angius, A. & Cas, A. *New Engl. J. Med.* **302**, 185-188 (1980).
- Little, P. F. R. *et al. Nature* **285**, 144-147 (1980).

- Caton, A. J. & Robertson, J. S. *Nucleic Acids Res.* **7**, 1445-1456 (1979).
- Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67**, 247-256 (1976).
- Choo, K. H., Cotton, R. G. H., Danks, D. M. & Jennings, I. G. *Biochem. J.* **181**, 285-294 (1979).
- Choo, K. H., Jennings, I. G. & Cotton, R. G. H. *Biochem. J.* **199**, 527-535 (1981).
- Huddleston, J. A. & Brownlee, G. G. *Nucleic Acids Res.* **10**, 1029-1038 (1981).
- Casadaban, M. J. & Cohen, S. N. *J. molec. Biol.* **138**, 179-207 (1980).
- Gergen, J. P., Stern, R. H. & Wensink, P. C. *Nucleic Acids Res.* **7**, 2115-2136 (1979).
- Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157-1174 (1978).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).

Intron-exon splice junctions map at protein surfaces

Charles S. Craik, Stephen Sprang, Robert Fletterick & William J. Rutter

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94142, USA

There have been several suggested explanations for the presence of noncoding intervening sequences in many eukaryotic structural genes. They may be examples of 'selfish DNA'^{1,2}, conferring little phenotypic advantage, or they may have some importance in gene expression and/or evolution. It has been suggested that each exon (coding sequence) may represent a structural or functional unit of the encoded protein^{3,4}, for which there is good evidence in the case of immunoglobulin⁷ and haemoglobin^{8,9} genes. Exon modification, duplication and recombination may thus be general mechanisms for the rapid evolution of eukaryotic structural genes. In many cases, however, it is not apparent that an exon corresponds to some specific feature of the encoded protein^{10,11}. We describe here evidence that intron-exon junctions usually map to amino acid residues located at the protein surface, suggesting a restriction on the permitted positions of introns within a gene.

For several proteins, proteolysis cleavage points are observed at the region of the splice junctions. This implies that the junctions correspond to a location at the protein surface. For example, clostripain specifically cleaves the three arginines which exist at the splice junctions, whereas another arginine, presumably buried, is less labile⁸. Similarly, the amino acids that are cleaved in the maturation of chymotrypsin correspond

to one of the splice sites (see Table 1). In addition, in immunoglobulins and proinsulin, proteases recognize surface amino acids mapping at or near splice junctions. We therefore have examined those proteins for which both genomic sequence and protein tertiary structure are known. Unfortunately the number of examples for which these data are available is limited. We have listed in Table 1 all those proteins that we know of for which both the gene structure and tertiary structure have been determined. This data set contains 31 splice junctions for 10 protein chains containing 1,928 amino acids. For all the splice junctions in Table 1, except one (the exon A boundary of lysozyme), the splice junctions map at the surface of the protein structure (for stereo views of dihydrofolate reductase and chymotrypsinogen, see Fig. 1).

We have quantified this observation by comparing the solvent-accessible surface area of amino acids encoded at intron-exon boundaries with those located elsewhere in the protein, a residue with $\leq 20 \text{ \AA}^2$ accessible surface being defined as interior¹². Note that this calculation does not always distinguish residues at the protein surface from those within its interior. For example, in a β -sheet core or inward facing surface of a helix, the immediate environment of a residue could be exposed to a solvent channel, or a surface amino acid can be buried by the side chains of neighbouring peptides. The calculation¹³ of the solvent-accessible surface for the amino acids that map at the splice junction shows that in every case except for lysozyme, at least one and usually both amino acids that map at the splice junctions are solvent-accessible.

A statistical analysis of these data can be made by assuming we have a random sample of events for which there are two possible states; that is, the splice junction maps either at the surface or at the interior of the protein molecule. Furthermore, the junction may fall either within (class I) or between (class II) codons. In the first case, accounting for 9 of the 31 samples, the intron-exon boundary maps to a single amino acid. The solvent-accessible surface calculation was used to find the fraction of buried amino acids (Table 1, column 6). Of the 1,928

amino acids represented, 745 are not accessible to the solvent. Thus for this sample, on average 0.389 of the amino acids are buried. Class II (the remaining 22) splice junctions map between two residues, thus the accessibility of both must be considered. Here position is determined to be interior only if both residues are inaccessible. Of the 1,918 adjacent pairs in the data set, 381 or 0.199, are buried by this criterion. However, none of the pairs at intron-exon boundaries are buried. Assuming that a binomial distribution describes the probability of a splice junction mapping to the protein interior, we can calculate the expected probability for the event that up to 1 out of 9 class I, and 0 out of 22 class II junctions, are buried by chance using the formula:

$$P = \sum_{y=0}^1 \binom{9}{y} p_1^y q_1^{9-y} \binom{22}{0} p_{11}^0 q_{11}^{22}$$

where $p_1 = 0.389$, $p_{11} = 0.199$ and q_1 , the probability of the junction being on the surface, is $1 - p_1$. This calculation shows that the probability P of observing only one (or zero) buried splice site for these samples is 6×10^{-4} . This calculation strongly suggests that the observed surface location of splice junctions is not coincidental.

The above observations might be explained by the restricted nucleotide composition observed around the splice junction¹⁴. This could result in a skewed amino acid composition that, in turn, might influence the location of this site within the tertiary structure of the protein. Thus, if charged amino acids were favoured there would be a 90% probability that they appear on the protein surface¹⁵. Similarly, as two-thirds of the amino acids in the hydrophobic core of proteins are Phe, Cys, Val, Ile, Leu or Met a choice from this set would tend to yield an interior splice junction. RNA sequence analysis¹² has shown that the sequence GT is nearly always found at the 5' junction, and AG at the 3' junction. We have translated in all reading frames the nucleotide sequence at the 5' and 3' side of all known splice junctions (refs 16, 17, Table 1 and unpublished data). The results show no particular constraints on the amino acid

Table 1 Solvent-accessible surface areas for intron-exon junctions in coding regions

Protein*	Intron	Intron-exon* splice position	Amino acid position	Amino acid solvent-accessibility ($\text{\AA}^2$)	Fraction of residues buried	Fraction of residue pairs buried
Cytochrome <i>c</i> ¹⁸		Gly	56	66	0.330	0.127
Immunoglobulin ¹⁹	A	Arg	103	150		
Fab λ light						
Immunoglobulin ¹⁹	A	Ser Gly	384 385†	90 46	0.356	0.183
Fab γ heavy	B	Glu Pro	423 424†	96 44		
Haemoglobin α ²⁰	A	Glu Arg	30 31	43 81	0.326	0.121
	B	Lys Leu	99 100	142 17		
Haemoglobin β ²⁰	A	Arg Leu	30 31	83 11	0.294	0.110
	B	Arg Leu	104 105	135 23		
Lysozyme ²¹	A	Trp	28	18	0.370	0.188
	B	Ser Ala	81 82†	74 33		
	C	Trp Val	108 109†	11 86		
Trypsinogen ²²	B	Ser	59	85	0.470	0.237
	C	Thr	144	91		
	D	Gln Gly	189 190	137 13		
Chymotrypsinogen B ²³	B	Gln Asp	34 35	25 38	0.443	0.278
	C	Thr	61	62		
	D	Lys Val	87 88	110 25		
	E	Ala Asn	146 147	126 48		
	F	Met Gly	192 193	105 24		
Dihydrofolate reductase ²⁴	A	Arg	28	159	0.338	0.151
	B	Gly	45	51		
	C	Lys	80	171		
	D	Lys Ala	122 123	114 54		
	E	Glu Tyr	161 162	105 32		
Carboxypeptidase A ^{25§}	C	Asp Glu	16 17	87 59	0.464	0.278
	D	Asn Arg	58 59†	68 38		
	E	Lys Lys	84 85†	15 36		
	F	Trp Arg	126 127†	36 43		
	G	Ala Ser	156 157†	23 30		
	H	Lys Thr	216 217†	86 108		
	I	Asn Gln	260 261	62 54		

* The amino acid residues represented are those of the protein crystal structure.

† Where the intron-exon boundary splits a codon a single residue is shown; otherwise residues on both sides of the junction are listed (see below).

‡ The precise position of the splice is sometimes equivocal due to the ambiguities of the DNA sequence around the intron-exon boundary so two amino acids are shown.

§ Five of the seven splice junctions were determined by detailed restriction and heteroduplex mapping (ref. 11 and C.S.C., unpublished).

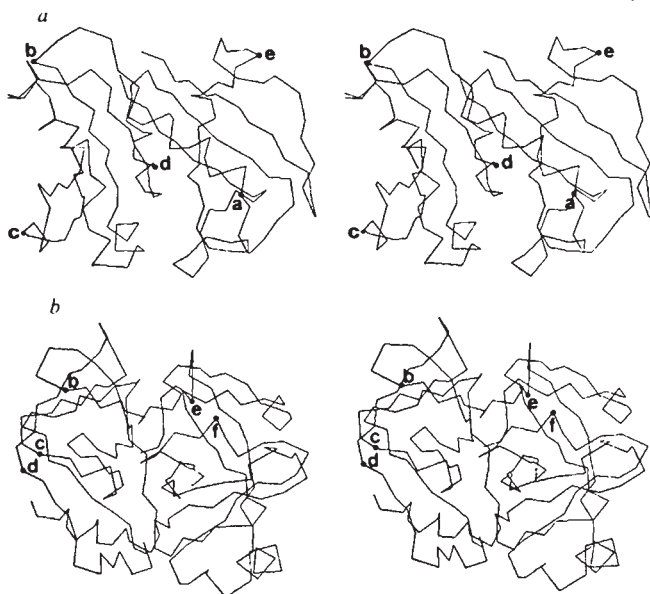


Fig. 1 Stereo views of the α -carbon chain tracing of dihydrofolate reductase (a) and chymotrypsinogen (b) showing the location of the splice junctions mapped on the protein sequence.

composition at that 3' side of the AG at the splice site. However, at the 5' side of the GT at the splice junctions, the extremely hydrophobic, and usually buried, amino acids (Phe, Cys, Val, Ile, Leu and Met) are underrepresented by a factor of 2.1 (expected from a natural abundance of 27%, that observed here being 13%) and the extreme hydrophilic and usually surface amino acids (Lys, Arg, Glu, Gln and Asp) are overrepresented by a factor of 1.6 (expected 28%, observed 44%). We have also tabulated the codons actually used at the junction sites. Again there was no strong bias in the amino acids present at the 3' side of the splice junction, but at the 5' side, the extremely hydrophobic amino acids are underrepresented by a factor of 4.1 (expected 27%, observed 6.6%) and the extremely hydrophilic amino acids are overrepresented by a factor of 2.3 (expected 28%, observed 64%). These data indicate that a hydrophilic amino acid is indeed more common near the intron-exon splice junction and provides a bias for the location of intron junctions at protein surfaces.

It is possible that the relationship between the location of the splice junction in the gene at the surface of the protein confers a biological advantage and hence is a result of natural selection. Introns and their associated splicing systems could be exploited in many ways during the evolution of a protein. We now realize that introns do not always mark large domains of protein structure, but often mark subdomains or segments of protein structure with some function. A further generalization can be made about the location of introns from the data reported here—usually the intron positions map outside the hydrophobic core and near the surface of the proteins. The cytoplasmic proteins used here are presumably exposed to a hydrophilic environment, thus the amino acids in the region of surface splice junctions are largely hydrophilic. If the rule is general, then proteins exposed to a hydrophobic environment should have hydrophobic amino acids near the surface splice junctions. That the 'seam' of a protein should appear on its surface is not surprising from an engineering viewpoint; if we consider the consequences of joining protein segments by cavalier recombination of exons, any novel amino acid sequences that might result from the new splicing would be isolated on the surface of the protein. This provides a means for change without forfeiting the pre-existing three-dimensional structure. It might therefore be expected that protein sequences contain additions or deletions in the region encompassing the splice junction. We have shown elsewhere that the region of major variation in the serine protease family occurs at or near splice junctions¹⁷.

Our findings are in accord with previous speculations on function or nonfunction and placement of introns, but set a requirement for their location with respect to protein structure. If the surface rule for intron-exon junctions is general, then the location of splice points in genomic DNA sequences also provides useful tertiary structural information about the gene product, as certain regions of the primary sequence can be reliably predicted to reside at the surface. Thus these data might be used in part to predict aspects of tertiary structure as well as to provide a strategem for the design of new proteins.

We thank David Standing and Pat O'Farrell for discussions; Graeme Bell, Carmen Quinto, Qui Lim Choo and Robert T. Schimke for communicating their sequencing data on chymotrypsin, trypsin and dihydrofolate reductase before publication; Raymond MacDonald for providing the trypsin cDNA; and Leslie Spector for preparation of the manuscript. This research was supported by NSF grant PCM 8001950 (to R.J.F.) and NIH grant AM 21344 (to W.J.R.). C.S.C. is a recipient of an American Cancer Society postdoctoral fellowship.

Received 21 January; accepted 13 July 1982

1. Doolittle, W. F. & Sapienza, C. *Nature* **284**, 601–603 (1980).
2. Orgel, L. S. & Crick, F. H. C. *Nature* **284**, 604–607 (1980).
3. Gilbert, W. *Nature* **271**, 501 (1978).
4. Blake, C. C. F. *Nature* **273**, 267 (1978); **277**, 598 (1979).
5. Toneyawa, S., Maxam, A., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1485–1489 (1978).
6. Sakano, H. *et al.* *Nature* **277**, 627–633 (1979).
7. Bernard, O., Hozumi, N. & Toneyawa, S. *Cell* **15**, 1133–1144 (1978).
8. Craik, C. S., Buchman, S. R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1384–1388 (1980).
9. Craik, C. S., Buchman, S. R. & Beychok, S. *Nature* **291**, 87–90 (1981).
10. Stein, I. P., Catterall, J. F., Kristo, P., Means, A. R. & O'Malley, B. W. *Cell* **21**, 681–687 (1980).
11. Quinto, C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 31–35 (1982).
12. Janin, J. *Nature* **277**, 491–492 (1979).
13. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1981).
14. Breathnach, R., Benoist, C., O'Hara, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
15. Sprang, S., Yang, D. & Fletterick, R. J. *Nature* **280**, 333–335 (1979).
16. Breathnach, R. & Chambon, P. *Rev. Biochem.* **50**, 349–383 (1981).
17. Craik, C. S. *et al.* in *Gene Regulation* (O'Malley, B. & Fox, C. F. (1982); *UCLA Symp. molec. cell. Biol.* **26**, (in the press).
18. Swanson, R. *et al.*, *J. biol. Chem.* **252**, 759–776 (1977).
19. Amzel, L. M. & Poljak, R. J. *Rev. Biochem.* **48**, 961–997 (1979).
20. Ten Eyck, L. F. & Arnone, A. *J. molec. Biol.* **100**, 3–29 (1976).
21. Blake, C. C. F. *et al.*, *Nature* **206**, 757–763 (1965).
22. Chambers, J. L. & Stroud, R. M. *Cryst.* **B35**, 1861–1874 (1979).
23. Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. *Nature* **214**, 652–660 (1967).
24. Volz, K. *et al.* *J. biol. Chem.* (in the press).
25. Lipscomb, W. N. *Acc. chem. Res.* **3**, 81–89 (1970).

High-frequency genomic rearrangements involving archaeobacterial repeat sequence elements

Carmen Sapienza*, Michael R. Rose† & W. Ford Doolittle*

Departments of *Biochemistry and †Biology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H7

Halobacterium halobium is an obligately halophilic archaeobacterium¹ of interest to molecular biologists for many reasons^{2–5}, one of which is the unexplained high frequency (10^{-4} – 10^{-2} mutants per cell plated) at which it yields readily identifiable and unstable mutants⁴. We showed previously⁵ that the genome of *H. halobium* contains many (>50) families of repeated sequences whose members are dispersed on both chromosome and plasmid. Here we report that most if not all of the members of most of these repeat sequence families effect or are affected by spontaneous genomic rearrangements. Quantitative analyses show that such repeat sequence-associated rearrangements (which may be of several kinds) occur at high frequencies ($>4 \times 10^{-3}$ events per family per cell generation), while unique-sequence DNAs are physically stable. The presence of so many families of elements of such great instability in the halobacterial genome gives it an unusual degree of structural and perhaps functional plasticity.

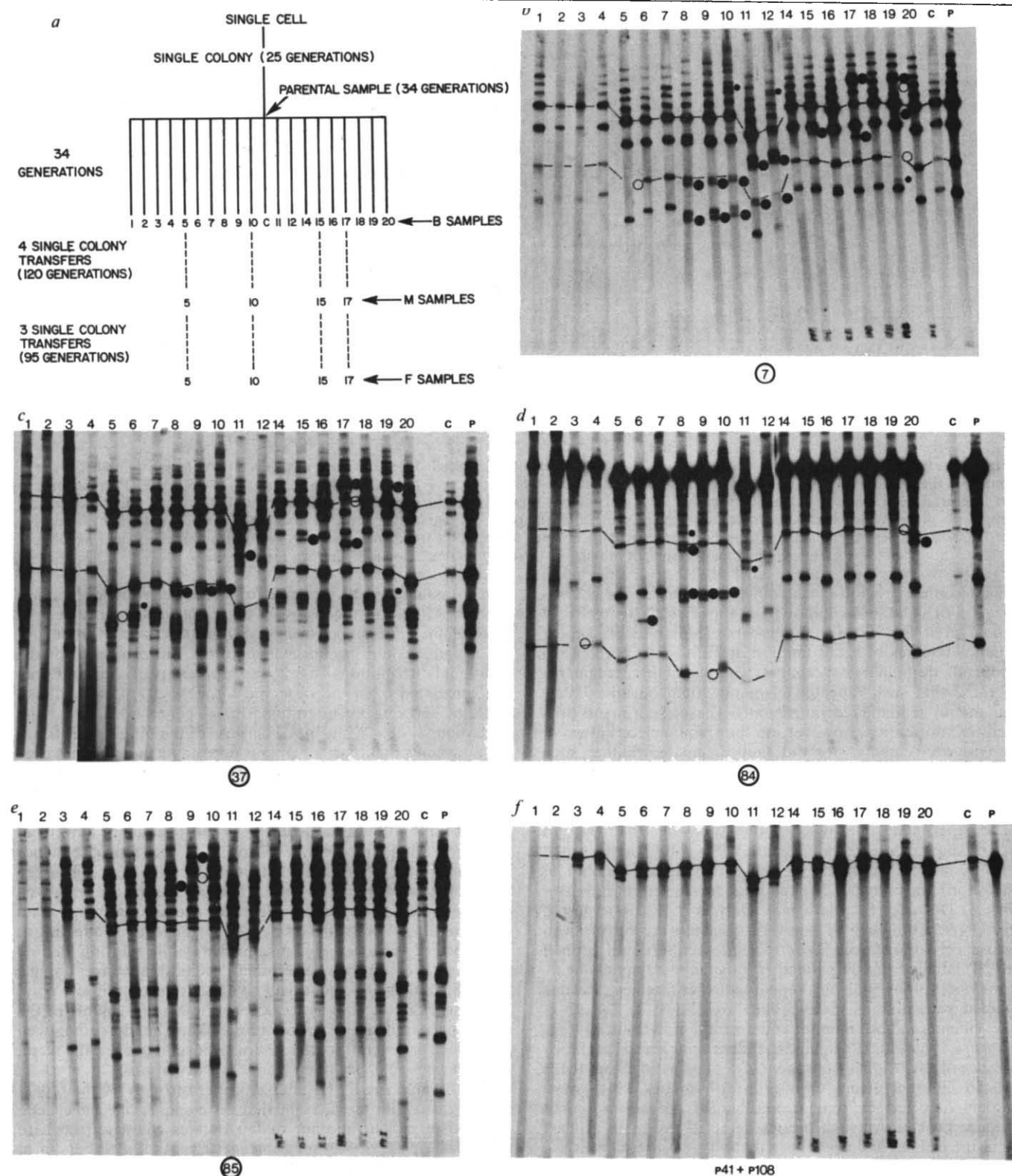


Fig. 1 *a*, Experimental protocol. A single cell of *H. halobium* wild-type strain NRC-1 was plated to give a single colony (25 generations), and cultured in liquid for a further nine generations. Most of this culture was used to prepare DNA taken to be representative of the first 'parental' cell (tracks labelled P in *b-f*). (A portion of the culture was diluted and grown in liquid for a further 34 generations, without re-plating. DNA prepared from this subculture appears in tracks labelled C of *b-f*). The rest of the culture was plated out to give single colonies; 19 of these which showed no visible phenotypic alteration were picked into liquid and grown for DNA preparation (tracks 1-19 in *b-f*). Again, 34 generations elapsed between initial plating and collection. Then four isolates (5, 10, 15 and 17) were re-plated and carried through seven additional platings and single-colony isolations. DNA was prepared from samples after four platings (120 generations, tracks labelled M in Fig. 2) and again after seven platings (215 generations, tracks labelled F in Fig. 2). The methods used for culture and DNA isolation are described in ref. 5. *b*, Results of probing *Eco*RI digestion products of the DNAs described above with cloned *H. halobium* repeat sequence element 7. ●, Hybridizing fragments present in the DNA of one of the 19 isolates which were absent from the DNA of the initial cell from which they derive (represented by tracks labelled P). ○, Hybridizing fragments absent from the DNA of one of the 19 isolates but present in the DNA of the initial cell. Hybridizations were performed as described previously⁵. *c-e*, Results for cloned repeat sequences 37, 84 and 85, respectively. *f*, Results for two cloned *Pst*I fragments of *H. halobium* DNA bearing only unique sequences.

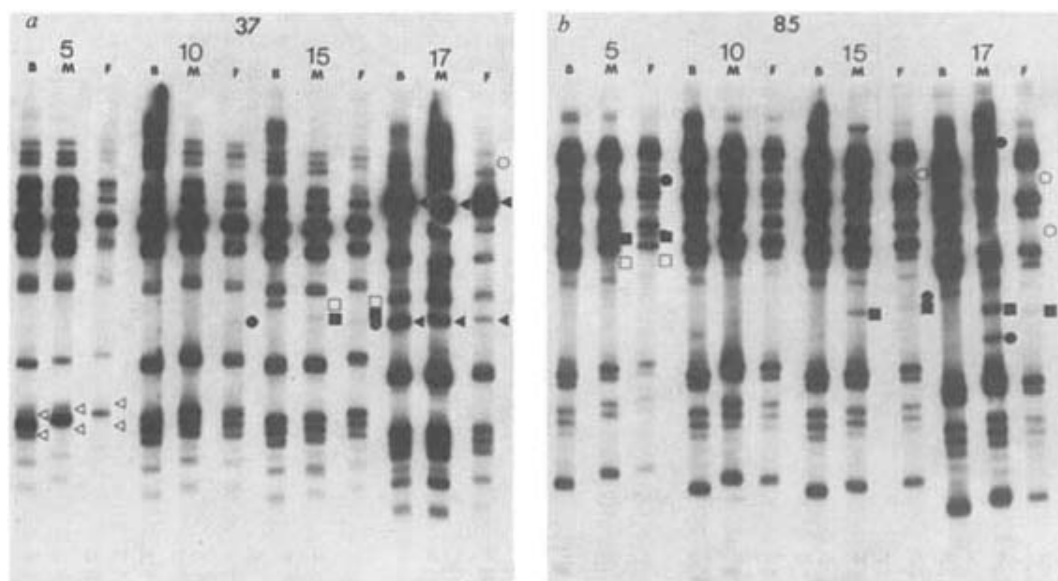


Fig. 2 *Eco*RI-cleaved DNAs from isolates 5, 10, 15 and 17 probed with 32 P-labelled cloned repeat sequence DNAs from strain NRC-1. ◀ Indicates the position of a fragment absent from the parental DNA sample (lane P in Fig. 1b–f) but present in B, M and F DNA samples. ◁ Indicates the absence of fragments from the B, M and F DNA samples that were present in the parental DNA sample. ■, Hybridizing fragments in the M and F DNA samples not present in B DNA samples; □, fragments hybridized in B DNA samples but absent from M or F DNA samples. ●, Fragments hybridized in only M or F DNA samples; ○, fragments absent from F but present in B and M DNA samples. *a* and *b* are results of probing with repeat sequence clones 37 and 85, respectively.

We described previously⁵ the molecular cloning and characterization of small (~3 kilobase pair, kbp) *Eco*RI fragments of *H. halobium* DNA bearing members of one or more repeat sequence families of small family size (2 or 3 to perhaps 20 members). Preliminary experiments with two of these fragments suggested that such repeated elements could be affected by spontaneous genomic rearrangements. These experiments provided no information, however, on the frequency or nature of those rearrangements, nor did they elucidate whether such instability is associated only with repeated elements and, if so, whether all families, or all members of any family, exhibit instability.

The experiments shown schematically in Fig. 1*a* measured rates of spontaneous rearrangement in 19 cell lines separated from a common ancestor by only that minimum number of generations (in practice, 34) required for the preparation of sufficient DNA for Southern blot analyses^{5,6}. *Eco*RI-digested total DNAs from each of these 19 isolates were resolved on agarose gels, transferred to nitrocellulose filters and probed with 32 P-labelled, *Eco*RI-cloned *H. halobium* fragments bearing members of seven distinct repeat sequence families⁵. Results obtained with four of these probes are shown in Fig. 1*b–e*. Many changes in both number and position of hybridizing fragments are apparent. If each independent appearance or disappearance of a hybridizing fragment is defined as an event, there were 56 different events among the 19 isolates which were uniquely detected by one or another of the seven repeat sequence probes. All such probes detected events in at least one of the isolates and we believe that all repeat sequences are associated with genomic rearrangements occurring at high and roughly comparable frequencies. On the other hand, two unique-sequence DNA probes detected few if any changes in any of the isolates (Fig. 1*f*).

If genomic rearrangements involving repeat sequences occur independently, then the observed frequency of events (roughly 10^{-2} per repeat sequence family per cell generation) could be taken as the probability of such rearrangements occurring in any cell line (isolate). However, events were not randomly distributed among isolates. Four isolates together accounted for about half the events and one showed alterations with five of the seven repeat sequence probes, while seven isolates showed no changes with any probe. If we assume that one event is usually coupled with others, then the probability of that first

event (assumed to be the same for all repeat sequence families) provides the most conservative estimate of rearrangement frequency. This can be calculated from the observed frequency of isolates exhibiting no changes using the zero-term of the binomial distribution. That is, the probability of change, P , can be calculated from the observed frequency ($7/19 = 0.368$) of isolates showing no alterations in 34 generations, using the equation $0.368 = [(1 - P)^{34}]^7$. This use of the binomial distribution assumes that rearrangements depend on independent Bernoulli trials over seven families. If rearrangements can arise at any time, however, a similar calculation using the zero-term of the Poisson formula is more realistic⁷. Both calculations give values for P of 4×10^{-3} events per repeat sequence family per cell generation (or from 4×10^{-4} to 1×10^{-3} per repeat sequence element). This is not remarkably lower than 10^{-2} , our observed frequency irrespective of order, because events are only partially coupled. From these values, we predict that about one-tenth of all isolates will experience rearrangements affecting at least one of the seven repeat sequence families during the first three or four divisions after their establishment from single cells. Such rearrangements will give rise to restriction site polymorphisms detectable as weakly hybridizing fragments in the DNA prepared after 34 generations. Weak bands were indeed detected for several isolates and with several probes and were, as far as possible, excluded from our calculations.

Of the nine *Eco*RI fragments probed strongly by the repeat sequence family member carried on clone 7, five have been deleted in one or another of the isolates characterized here (Fig. 1*a*) or previously⁵. This suggests that all members of at least this repeat sequence family are equally liable to exhibit genomic rearrangement. To determine whether the 'new' hybridization patterns produced by rearrangement are as stable as 'old' ones, 4 of the 19 isolates (Fig. 1*a*) were carried through seven further single-colony isolations, with DNA prepared as above after the fourth (120 generations) and seventh (215 generations) platings. These DNAs were probed with five cloned *H. halobium* repeat sequence DNAs and two cloned repeat sequence DNAs from *Halobacterium volcanii* previously shown to be homologous with two different *H. halobium* repeats⁵ (Fig. 2 and data not shown). In none of the three cases in which one of these four isolates had lost a hybridizing fragment during the first 34 generations was that fragment regained during 215 subsequent generations; vacated sites

appear not to be selectively reoccupied. In only one of the four cases in which one of these four isolates had gained a new hybridizing fragment during the first 34 generations was that fragment lost during the subsequent 215; newly-occupied sites do not appear to be uniquely unstable. Other changes observed during the propagation of these four isolates occurred with a frequency ($\sim 2 \times 10^{-3}$) in reasonable agreement with that obtained from the more extensive survey reported above.

The events observed result in either: (1) the appearance of a new hybridizing fragment without the loss of pre-existing fragments; (2) the loss of a pre-existing fragment without the appearance of a new one; or (3) the appearance of a new fragment and the loss of a pre-existing fragment, at about the frequency expected for independent occurrences of events of the first two types in the same cell line.

There is substantial indirect qualitative evidence^{4,8} for the transfer of discrete DNA sequences between plasmid, phage and chromosomal DNAs and we believe that many of the events whose frequency we have measured here are probably at least formally analogous to transpositions, although we cannot exclude alternative interpretations invoking deletions and duplication via unequal cross-over events between repeats of the same family. Whatever the nature of these events, the high frequency with which they occur in all repeat sequence families tested, and the large number of such families, must give the halobacterial genome an unusual degree of structural plasticity. It is easy to calculate that two daughter cells produced by a single division have only an 80% chance of bearing identical genomes, and that of the 10^{11} progeny in a 1-l culture established from a single cell, only 3×10^7 will bear genomes identical to that of the first cell, if all rearrangements are selectively neutral. They are not likely to be, and questions about the functional significance of this genomic plasticity and its relationship to the observed high frequency of spontaneous mutation are of considerable interest.

We thank the MRC of Canada for support, A. Cohen for technical assistance and N. Kleckner for critical advice.

Received 17 May; accepted 8 July 1982.

1. Woese, C. R. *Scient. Am.* **244**(6), 94–106 (1981).
2. Bayley, S. T. & Morton, R. A. *CRC Crit. Rev. Microbiol.* **6**, 151–205 (1978).
3. Weidinger, G., Klotz, G. & Goebel, W. *Plasmid* **2**, 377–386 (1979).
4. Pfeifer, F., Weidinger, G. & Goebel, W. *J. Bact.* **145**, 375–381 (1981).
5. Sapienza, C. & Doolittle, W. F. *Nature* **295**, 384–389 (1982).
6. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
7. Feller, W. *An Introduction to Probability Theory and its Applications* 3rd edn (Wiley, New York, 1968).
8. Schnabel, H. *et al.* *EMBO J.* (in the press).

The helicity of DNA in complexes with RecA protein

Andrzej Stasiak & Elisabeth Di Capua

Institute for Cell Biology, Swiss Federal Institute of Technology, Hoenggerberg, CH-8093 Zürich, Switzerland

The RecA protein of *Escherichia coli* is involved in recombination (for review see ref. 1). The protein binds transiently to double-stranded DNA in the presence of ATP. In the presence of ATP γ S, a non-hydrolysable analogue of ATP, recA-DNA complexes are stable^{2,3}. Duplex DNA in these complexes is stretched by a factor 1.5 (ref. 4), and the complexes appear in the electron microscope as helical filaments with a pitch of ~ 100 Å and 6.2 recA units per turn covering 18.6 base pairs (bp)⁵. RecA crystals have a space group of similar helical parameters⁶. In order to understand the function of recA, it is necessary to describe the conformation of the DNA in the recA complex. Using a topological method, the present work determines the helicity of DNA in the complex. We find that the DNA helix follows the protein helix visible in the electron microscope and has 18.6 bp per turn, which corresponds to an unwinding of the DNA double helix by 15° per bp.

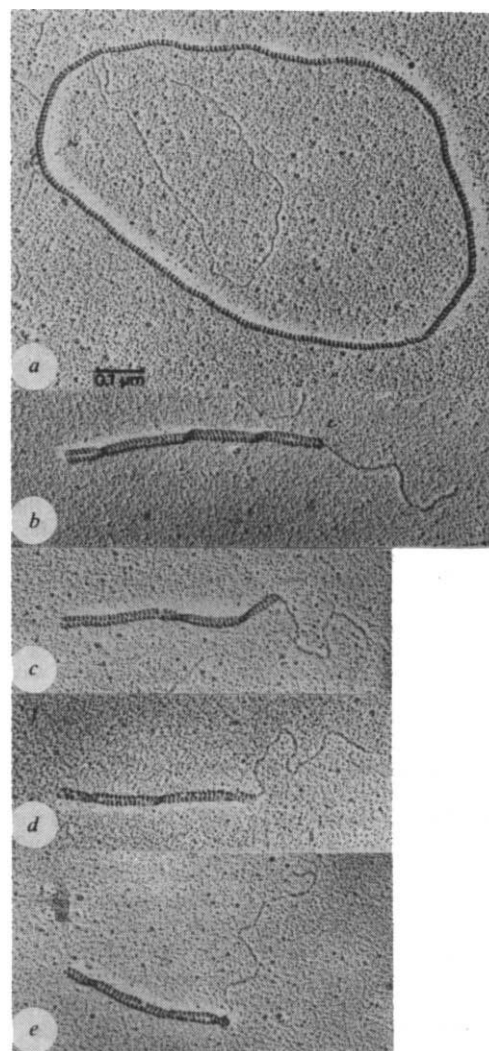


Fig. 1 Nicked circular DNA fully covered with RecA protein (a) and topoisomers of different linking numbers α reacted with RecA protein, $\alpha = 432$ (b), 445 (c), 459 (d), 472 (e). Note that in b to e the rod-like striated structures arise from side-by-side aggregation of neighbouring segments of the same circular RecA-DNA complex, while the thin filaments represent the highly positively supercoiled segments not covered by RecA. The cross-striations visible are due to the helical structure of the RecA-DNA complexes described in more detail in ref. 5. This helical structure is identical by our criteria whether the complexes are made with relaxed or cccDNA. The plasmid used is a pBR322 derivative of 4,961 bp. The topoisomers were produced by reacting plasmid DNA (100 μ M bp) with nicking-closing extract from rat liver nuclei¹³ in the presence of 0 μ M (b), 6 μ M (c), 12 μ M (d), 18 μ M (e) ethidium bromide. The recA-DNA complexes were obtained by incubating 50 ng DNA with 2.5 μ g RecA protein and 2 mM ATP at room temperature in 20 μ l 25 mM triethanolamine chloride (pH 7.6) and 5 mM magnesium acetate, conditions which seem to permit initiation of binding of RecA to cccDNA. After 20 min, ATP γ S to 0.5 mM was added to stabilize the complexes and the reaction was continued for 1 or 24 h at room temperature. The extent of reaction was the same after both times of reaction, thus implying that the reaction was complete after 1 h. Control experiments with different RecA-DNA ratios gave the same extent of DNA covering by RecA for a given DNA sample. Thus, in our experimental conditions, the extent of reaction is not a function of RecA:DNA ratio (see also ref. 4).

Double-stranded DNA has been shown⁷ to be unwound to some extent by RecA. We have now found conditions that allow the formation of complexes with covalently closed circular (ccc) DNA, and this enables us to determine the extent of this unwinding by a topological approach. A cccDNA molecule is characterized by its linking number which is fixed⁸; thus, reac-

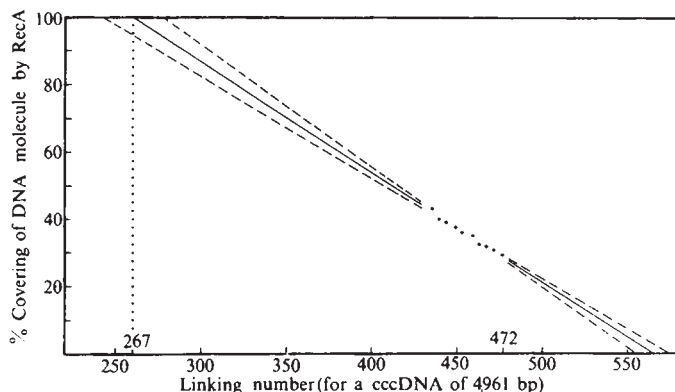


Fig. 2 Determination of the linking number of a cccDNA fully covered with RecA. The linking numbers were calculated by assuming 10.5 bp per turn of B-DNA in solution^{10,11}, which results in 472 turns for 4,961 bp, and subtracting from this value the number of superturns determined from the maximum of the distribution on band-counting gels⁹. For each topoisomer family the extent of covering with RecA was determined from 15–25 molecules. The standard deviation for each point was 2–3%, which is compatible with the presence of 7–8 topoisomers in each sample as seen on the band counting gels. Extrapolation to 100% covering with RecA leads to a value of 267 ± 19.2 for the linking number of a totally covered molecule. The confidence interval was drawn assuming 95% confidence.

tion with an unwinding agent will lead to a topological constraint in the molecule, and it is to be expected that the binding of the unwinding agent will stop when this constraint reaches a limit. The linking number will then be the sum of the underwound and overwound turns. In the case of RecA protein, which covers the DNA cooperatively⁴, we found (Fig. 1) that the underwound and the overwound regions are distinct and easily recognizable as striated complexes and highly supercoiled tails, respectively. Thus, the fraction of the molecule that is allowed to undergo complex formation can be measured. The linking number α is related to this complexed fraction by

$$\alpha = \frac{f_c n + f_i (N - n)}{360}$$

where n is the number of base pairs complexed with RecA, f_c the topological winding angle of the DNA complexed with RecA, N the total number of base pairs of the molecule and f_i the topological winding angle of the DNA not covered with RecA (recognized as a tail in Fig. 1).

If we now determine the extent of covering by RecA on cccDNA molecules of different α , the values of f_c and f_i can be determined by extrapolating the relationship to $n = N$ (completely reacted), and $n = 0$, respectively (Fig. 2). The angles f_c and f_i are assumed to be independent of α , that is, the constraint in the tail will be constant per unit length. We assume that this tail takes up a limit value of positive supercoils. Note that the values of f_c and f_i are the result of two orders of helicity, the primary helix of the DNA (in the covered part and in the naked tail respectively), and a secondary helix. In the case of f_i (uncovered tail) the second order helix is the positive supercoiling caused by torsional stress. In the case of f_c (complex) the second order helix is inherent in the structure of RecA–DNA complexes associated side by side (visible as turning-over points⁵) and is not caused by torsional stress; since this second-order helicity in the complexed part contributes only 2.6% to the linking number we do not take it into account here, and consider simply the overall unwinding.

cccDNA molecules of various linking numbers were obtained by reacting plasmid DNA with nicking-closing enzyme in the presence of various amounts of ethidium bromide. After purification of the products, the amount of unwinding was determined by estimating the supercoiling on band-counting gels⁹. Unfortunately, using the band-counting technique it is

difficult to determine precisely the linking number of highly supercoiled molecules above a superhelical density of 0.1. This limited the range in which the linking number and the extent of RecA covering could be related precisely. A series of cccDNA molecules of various linking numbers complexed with RecA protein is shown in Fig. 1; a relaxed complex formed with nicked DNA (thus allowing 100% complex formation) is shown for comparison. It is seen that the lower the linking number, the more RecA is bound. In Fig. 2, the extent of covering with RecA is plotted against the linking number. Least-squares analysis fits the data to a straight line with a correlation factor of 0.99. Extrapolation shows that 100% covering of DNA by RecA would require a linking number of 267 ± 19.2 (confidence 0.95). From our previous work, we know that a relaxed circular molecule of 4,961 bp will form a complex with 267 or 268 visible helix turns (ref. 5, and Fig. 1a). The linking number calculated above from Fig. 2 is very close to this number. We thus conclude that the RecA–helix seen in the electron microscope reflects the helix of the DNA molecule within the complex.

It is interesting to point out that the RecA self-complexes (that is, filaments formed in the absence of any DNA) form a similar helix as far as we can judge from our electron micrographs (see also ref. 2). By interaction with RecA protein, double-stranded DNA is unwound from about 10.5 bp per turn^{10,11} to 18.6 bp per turn. This corresponds to an unwinding of 15° per bp, a value similar to the unwinding by ethidium bromide (26° per 2 bp¹²) and is thus compatible with the intercalation type of interaction we suggested to explain the stretching of DNA in the complex⁴. We cannot exclude, however, the possibility that the DNA in the complex is melted and the single strands are wound around each other without forming a base-stacked helix.

We thank Professor Theo Koller for his help and support, Philipp Bucher for help with the calculations and Dr Martin Gellert for stimulating comments. This work was supported by Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung (grant to Th. Koller).

Received 10 May; accepted 15 July 1982.

1. Radding, C. M. *Cell* **25**, 3–4 (1981).
2. McEntee, K., Weinstock, G. M. & Lehman, I. R. *J. biol. Chem.* **256**, 8835–8841 (1981).
3. West, S. C., Cassuto, E., Mursallim, J. & Howard-Flanders, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2569–2573 (1980).
4. Stasiak, A., Di Capua, E. & Koller, Th. *J. molec. Biol.* **151**, 557–564 (1981).
5. Di Capua, E., Engel, A., Stasiak, A. & Koller Th. *J. molec. Biol.* **157**, 87–103 (1982).
6. McKay, D. B., Steitz, T. A., Weber, I. T., West, S. C. & Howard-Flanders, P. *J. biol. Chem.* **255**, 6662 (1980).
7. Cunningham, R. P., Shibata, T., Das Gupta, C. & Radding, C. M. *Nature* **281**, 191–195 (1979).
8. Crick, F. H. C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2639–2643 (1976).
9. Keller, W. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4876–4880 (1975).
10. Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 200–203 (1979).
11. Rhodes, D. & Klug, A. *Nature* **286**, 573–578 (1980).
12. Baase, W. A. & Johnson, W. C. *Jr Nucleic Acids Res.* **6**, 797–814 (1979).
13. Champoux, J. J. & McConaughy, B. L. *Biochemistry* **15**, 4638–4632 (1976).

Errata

The cover photograph on the issue of 22 July, relating to the letter 'Monoclonal antibodies against human T lymphocytes label Purkinje cells of many species' by J. A. Garson *et al.* *Nature* **298**, 375–377 (1982), was reproduced in only two colours so that orange and green staining was not sufficiently differentiated. The photomicrograph shows neonatal mouse cerebellar cortex stained with two monoclonal antibodies simultaneously. The Purkinje cells are stained orange by Leu-4/biotin/avidin/rhodamine, while basket fibres and other axons are stained green (fluorescein conjugate) with the antineurofilament monoclonal RT97 (donated by B. Anderton).

In the article 'Dynamic observation of defect annealing in CdTe at lattice resolution' by R. Sinclair *et al.* *Nature* **298**, 127–131 (1982), parts *c* and *e* of Fig. 3 have been transposed thus interrupting the movement of the arrowed dislocation across the frame. The figure is correct on reprints.

MATTERS ARISING

Absence of thermal effects on photon mass measurements

THE mass of the photon^{1,2} is known to be less than 10^{-16} eV. However, in a letter to *Nature*, Primack and Sher³ have questioned the validity of this result. They suggest that electrodynamics may undergo a phase transition at low temperature, and point out that all of the photon mass experiments took place at temperatures above a few degrees Kelvin. Thus, they claim that the above mass limit may only apply to a high-temperature phase ($T > T_c$ where T_c could be as large as a couple of degrees Kelvin), and that when measured at lower temperatures the photon mass could be as large as about 10^{-4} eV. Apparently, low-temperature photon mass experiments are presently being considered⁴. We explain here why we believe that the scenario proposed by Primack and Sher³ is, in fact, impossible. We find that the limit of 10^{-16} eV on the mass of the photon is unaffected by the fact that the relevant experiments were performed at finite temperature. It is therefore completely valid and applies to experiments regardless of whether they are carried out at room temperature or at absolute zero.

To analyse thermal effects on measurements of the photon mass we must define what is meant by a phase transition in finite temperature field theory⁵⁻⁷. At zero temperature, the vacuum and the various particle states are described by definite state vectors. We can study vacuum expectation values of different field operators and we can measure the photon mass by examining how a photon propagates through the vacuum. At non-zero temperature, the system is described by a density matrix. We can no longer consider vacuum expectation values, but instead we deal with thermal expectation values. At finite temperature space is filled with thermal radiation and to measure the photon mass we must determine how a photon propagates through this thermal radiation. It is important to note that the vacuum does not change with temperature. However, quantities like field expectation values or particle masses can change because at non-zero temperatures they are measured not in the vacuum but in a thermal ensemble.

Therefore, we must ask what is the nature of the thermal fluctuations which occur at or below room temperature, and how can these affect the measured value of the photon mass? In particular, are the interactions of the photon with the thermal fluctuations at these temperatures strong enough to make a massive photon appear massless? The mass of a field sets a lower limit on the energy of its small fluctuations so a field of mass m will only experience significant fluctuations at tem-

peratures $kT \geq m$. At room temperature or below the only relevant thermal radiation consists of blackbody photons. Note that if there existed a fundamental or composite scalar field with a small vacuum expectation value which gave the photon a mass, then this field would only experience significant thermal fluctuations at normal temperatures if it produced a charged scalar particle with a mass of a fraction of an eV. If the associated charge scalar were heavier than this then thermal fluctuations of this scalar field would be exponentially suppressed. Since no such light charged particle exists we conclude that there is no way for thermal fluctuations to change the expectation value of such a field and thereby change the photon mass from some non-zero low-temperature value to zero at room temperature. However, we must still consider the effect of thermal photons on the photon mass at finite temperature.

The real-time thermal propagator for a particle of mass μ is

$$\frac{i}{k^2 - \mu^2} + \frac{2\pi}{e^{E/kT} - 1} \delta(k^2 - \mu^2) \quad (1)$$

The temperature-dependent term in this propagator is multiplied by $\delta(k^2 - \mu^2)$ so thermal effects are completely absent, to lowest order in α , for off-shell, virtual photons. The best limits on the photon mass come from measurements of the static magnetic field produced by Jupiter and these involve virtual photons. The second term in equation (1) is only relevant for real, on-shell photons of energies $E \leq kT$. For these photons this term represents the stimulated emission effect which the presence of thermal photons produces. Equation (1) is of course modified by order α corrections due to the interactions between thermal photons and a propagating virtual photon. However, such interactions occur only through loops of charged particles and are suppressed by a factor $e^{-m/kT}$ where m is the mass of the charged particle. The lightest charged particle is the electron and it produces polarization factors,

$$\Pi_{\mu\nu} \propto e^2 m_e^2 \left(\frac{m_e}{kT} \right) e^{-m_e/kT} \quad (2)$$

At temperatures of order room temperature or below, the factor $e^{-m_e/kT}$ is a gigantic suppression factor so these thermal effects are completely negligible. Thus, a virtual photon propagating through thermal radiation at normal experimental temperatures is completely unaffected by the presence of that radiation simply because no mechanism exists for it to interact in any appreciable way with the thermal photons. It follows that experiments which determine how a virtual photon propagates, like the photon mass experiments, will be unaffected as well. The photon mass limit of 10^{-16} eV is therefore valid at low temperature.

We wish to thank Alvaro DeRujula for

helpful discussions. This work was supported in part by the US Department of Energy under contract DE-AC02-76ER03230-A005.

L. F. ABBOTT

M. B. GAVELA*

*The Martin Fisher School of Physics,
Brandeis University,
Waltham,
Massachusetts 02254, USA*

* Permanent address: Laboratoire de Physique Theorique et Hautes Energies, University Paris-Sud, 91405 Orsay, France.

1. Goldhaber, A. S. & Nieto, M. N. *Rev. mod. Phys.* **43**, 277-296 (1971).
2. Davis, L., Goldhaber, A. S. & Nieto, M. N. *Phys. Rev. Lett.* **35**, 1402-1405 (1975).
3. Primack, J. R. & Sher, M. A. *Nature* **268**, 680-681 (1980).
4. Dombey, N. *Nature* **288**, 643-644 (1980).
5. Dolan, L. & Jackiw, R. *Phys. Rev. D* **9**, 3320-3341 (1974).
6. Weinberg, S. *Phys. Rev. D* **9**, 3357-3378 (1974).
7. Linde, A. D. *Prog. theor. Phys.* **42**, 389-435 (1979).

SHER AND PRIMACK REPLY—In our paper, we explained that neither spontaneous nor dynamical symmetry breaking could lead a mass for the photon at low temperature; but we speculated that some other physical mechanism could conceivably do it, and we pointed out that the question could be decided experimentally. Abbott and Gavela argue again that spontaneous symmetry breaking (the 'Higgs mechanism') will not work. The essential physical point they make is that "... a field of mass m will only experience significant fluctuations at temperatures $kT \geq m$." We agree. The question is whether the absence of charged particles (either bosons, mentioned by Abbott and Gavela, or fermions) lighter than the electron excludes an electromagnetic phase transition at $kT < m_e$, even one arising from a hitherto unknown mechanism. We are persuaded that their argument indeed excludes this possibility.

It should be noted that this argument could affect other, more realistic calculations. There have been many calculations of supercooling in both SU_3 and the Weinberg-Salam model. In some of these, the mass (at $T=0$) of the scalar before the transition vanishes (Coleman-Weinberg); the field can then experience significant fluctuations. In some, however, the scalar does have a small but non-zero mass m (see ref. 1). The argument of Abbott and Gavela appears to rule out a transition at $T \ll m$; completing a transition at such a temperature may not be possible (if only the Higgs potential is considered; massless fermions could condense and drive the transition if the coupling constant is large enough).

MARC A. SHER

JOEL R. PRIMACK

*Physics Department,
University of California,
Santa Cruz,
California 95064, USA*

1. Cook, G. P. & Mahanthappa, K. T. *Phys. Rev. D* **23**, 1321-1336 (1981).

Crystal instability and melting

COTTERILL AND MADSEN¹ argue against an essential connection between thermodynamic melting and crystal instability. Their molecular dynamics simulation of a 336-particle Lennard-Jones (LJ) assembly revealed the existence of a dislocation-mediated mechanical instability in the superheated crystalline phase at a density ($\rho_{\text{inst}} = 1.22$) lying just outside the melting interval ($\rho_{\text{cryst}} = 1.38$, $\rho_{\text{liq}} = 1.24$) determined by Hansen and Verlet² for the same isotherm. Moreover, with increasing temperature, the instability was progressively further removed from the melting interval. We suggest, however, that contrary to their hypothesis, these results add further to an already substantial weight of evidence for the essential connection between crystal instability and melting^{3,4}.

Born⁵ was among the first to propose that melting occurs as a mechanical instability when one of the crystal shear moduli falls to zero. Of course, experiment demonstrated that the shear moduli remained finite up to the melting point, but we have shown^{6,7} that extrapolation of one of the shear moduli as a function of density beyond the melting point, reveals that it would fall to zero at the freezing density for a wide range of ionic, metallic and molecular crystals. We interpreted this as implying the universal existence of an instability concealed between the melting and freezing lines but lying close to the latter. This is confirmed by theoretical calculations for alkali halides⁴ and, indeed, is what one would expect from a proper interpretation of Born's melting criterion. If the crystal were superheated to the critical density, ρ_c , of vanishing static shear modulus, the resultant localization of shear modes would contribute an additional amount, $R \ln 2$, to the entropy⁸, so that the system would be able thermodynamically to make an isothermal jump from a lower density, ρ_{cryst} to the density, ρ_c . Thus the instability, in normal circumstances, would be concealed by the first-order transition, and to a first approximation ρ_c ($= \rho_{\text{inst}}$) may be equated to ρ_{liq} .

The molecular dynamics result of Cotterill and Madsen seems to confirm our original interpretation of Born's melting criterion, since the freezing density and instability density are almost coincident. Both the suppression of long-wavelength phonons in their finite (336-particle) system and critical slowing would serve to delay the instability and thus it lies a little beyond the melting interval, and is displaced further away with increasing temperature.

Note also that the single-cell occupancy system simulated by Hansen and Verlet to obtain a reversible path between liquid and solid, probably has a first-order transition (L. V. Woodcock, personal communication) rather than the second-

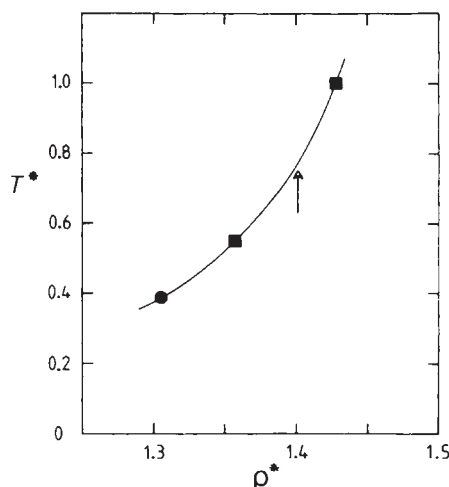


Fig. 1 LJ glass transition temperature (T^*) as a function of density (ρ^*). The expected transition temperature on Cotterill and Madsen's isotherm is indicated by the arrow.

order transition assumed by Hansen and Verlet. Their phase boundaries may not therefore be altogether reliable. We note that the actual freezing line for argon is displaced progressively away from the Hansen and Verlet line in the same manner as the instability line.

Finally, Cotterill and Madsen's location of the glass transition between the melting and freezing lines at $\rho_g = 1.29$ seems incorrect. Their procedure of finding the intersection of extrapolated linear fits of the extremes of their isothermal pressure-density data ignores the fact that the isotherms are naturally curved in any one phase. Figure 1 shows LJ glass transition temperatures as a function of density reported by Abraham⁹ (squares) and Damgaard-Kristensen¹⁰ (circle) and this predicts a glass transition on Cotterill and Madsen's 0.77 isotherm at $\rho_g = 1.401$; that is, well above their value $\rho_g = 1.29$, and outside the melting interval. That their glassy pair and bond angle distribution functions (pdf and adf, respectively) are actually evaluated for supercooled liquid is evidenced by the lack of a fully developed split-second peak in the pdf. Presumably the incipient splitting in their adf at 110° and 120° would be fully developed in a true glass.

J. L. TALLON

Physics and Engineering Laboratory,
DSIR, Private Bag, Lower Hutt,
New Zealand

1. Cotterill, R. M. J. & Madsen, J. U. *Nature* **288**, 467 (1980).
2. Hansen, J.-P. & Verlet, L. *Phys. Rev.* **184**, 151 (1969).
3. Jelinek, G. E. *Phys. Rev.* **B3**, 2724 (1971).
4. Boyer, L. L. *Phys. Rev. Lett.* **45**, 1858 (1980).
5. Born, M. *J. chem. Phys.* **7**, 591 (1939).
6. Tallon, J. L., Robinson, W. H. & Smedley, S. I. *Nature* **266**, 337 (1977).
7. Tallon, J. L. *Phil. Mag.* **39A**, 151 (1979).
8. Tallon, J. L. *Physics and Engineering Laboratory Rep. No. 705* (1981); *Phys. Lett.* **87A**, 361 (1982).
9. Abraham, F. F. *J. chem. Phys.* **72**, 359 (1980).
10. Damgaard-Kristensen, W. *J. Non-Cryst. Solids* **21**, 303 (1976).

COTTERILL AND MADSEN REPLY—Tallon considers our instability density for $T = 0.75$ to be indistinguishable from the freezing density, within computational error, and dismisses the disparate temperature dependencies with a reference to real argon, despite its differences from the pseudoinfinite Lennard-Jones matter studied by Hansen and Verlet¹ and ourselves². Moreover, the discrepancy between the two materials is opposite to Tallon's claim. We cannot see why model-dependent stabilization, beyond the instability point, should be more effective at higher temperature. Melting is a reversible first-order transition, the location of which depends on the properties of both liquid and crystal. Instability theories attempt comprehensive treatment of the transition, including projections of liquid properties, using exclusively crystalline characteristics. Such unilateral approaches are inferior to those incorporating more realistic concepts of liquid structure.

Because the properties of all three condensed states are nevertheless determined by the same interatomic interactions, empirical correlations between properties are inevitable; the proximity of the instability and freezing densities²⁻⁴ is thus hardly surprising. But vacancy⁵ and dislocation⁶ parameters both also correlate with the melting temperature, even though melting theories based on these two defects are mutually exclusive⁷. We show elsewhere⁸ that the excess entropy of liquids over crystals is associated with fluidity, not disorder, and stems from conversion of transverse phonons to diffusive modes⁹. Tallon's rather nebulous analysis¹⁰ fails through, amongst other things, neglecting the entropy associated with perfect crystal oscillatory modes. The location of the glass transition density is subject to a choice of convention, but a density of 1.4 for $T = 0.77$ is irreconcilable with our findings; the pressure varies linearly with density in the range 1.35–1.45. Applying Abraham's convention¹¹ to our data, or vice versa, produces no contradiction.

R. M. J. COTTERILL
J. U. MADSEN

Department of Structural
Properties of Materials,
The Technical University of Denmark,
DK-2800 Lyngby, Denmark

1. Hansen, J. P. & Verlet, L. *Phys. Rev.* **184**, 151 (1969).
2. Cotterill, R. M. J. & Madsen, J. U. *Nature* **288**, 467 (1980).
3. Ross, M. & Alder, B. J. *Phys. Rev. Lett.* **16**, 1077 (1966).
4. Hoover, W. G. & Ree, F. H. *J. chem. Phys.* **49**, 3609 (1968).
5. Gorecki, T. Z. *Metallkunde* **65**, 426 (1974).
6. Cotterill, R. M. J., Jensen, E. J., Damgaard-Kristensen, W., Paetsch, R. & Esbjørn, P. O. *J. Phys.* **36**, 35 (1975).
7. Cahn, R. W. *Nature* **273**, 491 (1978).
8. Madsen, J. U. & Cotterill, R. M. J. *Phys. Lett.* **83A**, 219 (1981).
9. Madsen, J. U. & Cotterill, R. M. J. *Physica Scripta* **24**, 959 (1981).
10. Tallon, J. L. *Phys. Lett.* **87A**, 361 (1982).
11. Abraham, F. F. *J. chem. Phys.* **72**, 359 (1980).

BOOK REVIEWS

A disappearing breed

Brian Bertram

WITH the publication of Vols IIIC and IIID of *East African Mammals*, Jonathan Kingdon has completed his *tour de force*. These final volumes deal with most of the artiodactyls — the bovines and antelopes whose numbers and diversity contribute so much to the array of East Africa's mammals.

The format is the same as in the earlier volumes. For each tribe the author presents an introductory profile covering topics such as structure, coat colouring, the form of the horns, body size, evolutionary history, signalling and ecological niches. For each species he then describes its past and present distribution, its distinguishing features, its feeding and predators, its behaviour and breeding. Again, the author's superb illustrations abound, and contribute greatly to the information presented; Kingdon can convey so much more with a few pencil lines than anyone could say in many pages of text. There are always beautiful and meticulously accurate drawings of the whole animal, and there are often equally accurate drawings of the skinned animal, revealing the underlying musculature, and of the skeleton. In addition, there are countless thumbnail sketches which capture the essence of an animal's appearance or behaviour.

Unlike earlier volumes in the series, neither of these final volumes can stand on its own. Volume IIIC includes the bovines and the smaller antelopes, while Vol. IIID has the larger antelopes, four appendices, the bibliography and indexes for these last two volumes, and a species checklist for all seven volumes. The first appendix outlines the fundamental problems of conservation, and the ethic, the threats, and the meagreness of the response. The second, by David M. Jones, summarizes current knowledge of the feeding and handling of mammals in captivity. Appendix III lists the various parasites and diseases known to infect

East African Mammals: An Atlas of Evolution in Africa, Vols IIIC and IIID. By Jonathan Kingdon. Vol. IIIC, pp.393, ISBN 0-12-408344-7; Vol. IIID, pp.353, ISBN 0-12-408345-5. (Academic: 1982.) Each volume £49.95, \$99.50.

East African mammals, and Appendix IV considers briefly the potential for commercial exploitation of wildlife.

We can now look back on the whole series, and we must do so with considerable awe. The 2,900 pages contain a vast wealth of information and discussion. They also contain hundreds of fascinating and often most beautiful illustrations. Particularly in the earlier volumes, more guidance could have been given to help the reader to make use of the illustrations, noting important features or explaining what an animal was doing. But it remains indisputable that these seven books constitute a remarkable work.

The series has expanded in the writing. "This is the first of a three volume work" announced the flysheet to Vol. I, in 1971. By geometrical progression, this was followed by IIA and IIB, and by IIIA, IIIB, IIIC and IIID. It has been a somewhat unplanned growth, resulting in some awkward anomalies of arrangement: thus the page numbers in Vols IIB and IIID, but only these, continue from where the previous volumes' numbers left off; the bibliographies are sometimes split up according to subject and sometimes not; and what is essentially the same gazetteer appears in Vols I, IIB, IIIA and IIID.

Expansion plus inflation have had an alarming effect. On publication eleven years ago, the 446 pages of Vol. I cost £12 (about £38 at today's prices), and a reprint can still be bought for £34.40. The 353 pages of Vol. IIID will cost you £49.95, and the whole series £255 (\$534.50), which

means that what started as a set of three, which keen private naturalists could just hope to afford, has ended as a series mainly for libraries. It is well worth making sure you have access to it somehow, because it contains so much of value as well as of beauty.

Valuable beautiful things are rare, and so may well prove to be good investments. On their current form these books will be. But I fear that the series, like its author and like several of the species it describes, is a disappearing breed. The very high standard of production, the generous use of space, the large typeface and the ample illustrations all allow what have become increasingly reference works for the specialist and the professional to retain the attractiveness of coffee table books. What a pity it is that other continents and other faunas have yet to be so well served by anyone as East Africa has by Jonathan Kingdon and Academic Press. □

The art of Jonathan Kingdon — a Greater Kudu from Vol. IIIC of *East African Mammals*. A limited edition of full-size illustrations from the series has also been published. Details can be obtained by writing to Pinguin Prints, PO Box, Oron, UK, enclosing a stamped, addressed envelope.



Brian Bertram is Curator of Mammals at the Zoological Society of London.

And citation begat co-citation . . .

A.C. Allison

ISI Atlas of Science: Biochemistry and Molecular Biology 1978/80. Pp.540. ISBN 0-941708-00-4. (Institute for Scientific Information, Philadelphia: 1982.) \$45 (individual), \$90 (institution).

THE information explosion is daunting to most people but challenging to others, including Eugene Garfield and his colleagues of the Institute for Scientific Information (ISI) in Philadelphia, who have devoted themselves to the systematic collection and organization of information. First came *Current Contents*, the usefulness of which is universally acknowledged. The commercial success of *Current Contents* allowed the production of the *Science Citation Index*, which is widely but less generally used. More recently the Institute has published *Research Fronts in ISI/Biomed* and now the *ISI Atlas of Science*, a reference work to 102 Research Fronts or "Specialties" within the fields of biochemistry and molecular biology.

The Research Fronts included in the *Atlas* were identified by frequent citations in the *Science Citation Index*. The central process of selection of papers was the process termed "co-citation clustering". This is when an author, in citing the works of two other authors, establishes an intellectual link between them. The papers quoted in the *Atlas* were both highly cited and frequently co-cited. This computer analysis is supposed to reveal sets of papers that are important, because of the attention they have received from colleagues of the authors, and highly related because of the co-citation linkage. For each specialty, maps are constructed in which the documents are arranged graphically according to a mathematical treatment of co-citation patterns. The same treatment is also used to construct a Global Map, which shows the relationships among the various specialties in biochemistry and molecular biology.

For each specialty there is a section entitled "Citation Core Documents" which comprises the main citations and co-citations constituting the cluster. This is intended to provide objective assessment of clustering and case documentation, instead of the usual subjective assessment by individual scientists. There is a short, one-page mini-review of each specialty.

The unique feature of the *Atlas* is the Specialty Maps and the Global Map, but it can be questioned whether this is the most useful way to employ computers. The value of computers for information retrieval, especially for recovering associated information (e.g. on prostaglandins in a particular cell type) is well established and has lightened this burden for most scientists. Listing the Citation Core Documents for each specialty is helpful,

although similar information can be obtained through the *Index Medicus* and other compilations. A brief perusal of these documents will often establish their relatedness more surely than the maps, which are for the most part statements of the obvious and are at times misleading.

It can also be questioned whether, for persons intending to go into a field seriously, reviews of specialties by various experts, such as are found in the *Annual Review of Biochemistry*, do not provide

equally accessible and more authoritative and up-to-date sources of information than the data in the *Atlas* (most of the analyses in the *Atlas* were completed before 1980, but supplements refer to major citations in 1980).

The general conclusion is that the thirst for information is not easily satisfied, and that the *Atlas* does provide a ready source of information on a range of subjects. Certainly, it is likely to be widely sold and used. □

Anthony Allison is Director of the Institute of Biological Sciences, Syntex Research, Palo Alto, California.

Mitochondria: the cell's paying guests?

M. J. Selwyn

Mitochondria. By Alexander Tzagoloff. Pp.342. Hbk ISBN 0-306-40799-X; pbk ISBN 0-306-40778-7. (Plenum: 1982.) Hbk \$42.50, £26.78; pbk \$19.95, £12.57.

MITOCHONDRIA have provided interesting and challenging subjects for study by biochemists and cell physiologists ever since it became possible, in the early 1950s, to isolate them in quantity and quality. The reason for this interest is that the organelles have an integrated set of functions which, although complex, is more simple than that of a whole cell; for instance it was soon recognized that they contained complete sets of enzymes for the citric acid cycle and fatty acid oxidation. In his new book, Alexander Tzagoloff appropriately describes these metabolic pathways in brief, in only one of the eleven chapters; most readers can be expected to be well acquainted with them.

The central chapters of *Mitochondria* deal with the properties and functions of the membrane-bound enzymes and carrier systems, including the vital function of regenerating ATP by oxidative phosphorylation. The story of the investigation of these systems should be a fascinating and salutary tale. The classical techniques and concepts of aqueous phase biochemistry failed to resolve the mechanisms, and new methods and ideas for membrane biochemistry had to be developed. Sad to say I found this, the longest and a very important section, somewhat disappointing. The author appears to give unenthusiastic acceptance to the chemiosmotic coupling theory and these chapters are not as up-to-date as they should be, few papers after 1978 being cited. The treatment is patchy and sometimes falls far short of what is needed. For example, descriptions of ions simply as "permeant" can be misleading, as is describing uncouplers as "blocking" reactions; and are we really not sure that

valinomycin is a mobile carrier and gramicidin a pore-former? The failure to describe the proton-conducting function of the membrane-embedded component detracts seriously from the treatment of the ATPase complex.

The last two chapters, on mitochondrial biogenesis and genetics, make more exciting reading. Over the past 20 years mitochondria have been found to contain not only their own DNA but a complete apparatus for synthesizing a few polypeptides. And whilst a decade ago the mitochondrial protein synthesizing system appeared to be very similar to that of prokaryotes, major differences are now being discovered — not only between the mitochondrial and prokaryotic systems but also between mitochondrial systems from different organisms. The idea that mitochondria are the descendants of symbiotic prokaryotes which paid for their keep in the ancestral eukaryote cell by regenerating ATP is now no more plausible than the alternative that mitochondria arose by segregation of certain functions and some DNA within the eukaryote. Whichever is the case, it is clear that mitochondria are now set to be a testing ground for fundamental investigations in molecular genetics and into the evolution of eukaryotic cells.

Tzagoloff helps to fill the gap left by the failure to revise Lehninger's *The Mitochondrion* since it first appeared in 1964. He provides much of the basic material which a research student working with mitochondria will need to know. Teachers of advanced undergraduate courses will also want a copy, and one or two in the library, but they will have to supplement and comment on the bioenergetics chapters. □

M. J. Selwyn is Dean and Senior Lecturer in Biochemistry in the School of Biological Sciences, University of East Anglia.

On the edges of the sea

John H. Steele

Ecology of Coastal Waters: A Systems Approach. By K. H. Mann. Pp.322. Hbk ISBN 0-632-00669-2; pbk ISBN 0-632-00953-5. (Blackwell Scientific/University of California Press: 1982.) Hbk £25, \$36; pbk £10.80, \$18.

COASTAL ecosystems are very diverse, reflecting the many different ways in which the land interacts with the sea. This coastal zone has always fascinated biologists and, more recently, has become subject to increasing pressure for recreational and commercial uses. Thus, for both scientific and practical reasons, we need to generalize about these environments. Dr Mann's book approaches the problem in two ways — by description of the principal types of coastal ecosystem and by considering the application to these systems of various modelling techniques.

Coastal waters are often considered as the most accessible parts of the pelagic ocean. Here only one chapter is given to phytoplankton-based systems, while most of the text is devoted to excellent analyses of major features such as marshes, mangroves and coral reefs. There is also an emphasis on the sediment communities and their microflora which play a critical role in exchange of materials across the sediment-water interface and from land to sea.

In particular, Mann focuses on general patterns of energy flow and on average conditions within each system. Although a chapter on water movements is included, there is little on the consequences for these coastal areas of the highly episodic nature of physical changes. Instead, the homeostatic properties of coastal ecosystems are a central feature of the approach used in this book. As Mann acknowledges, further work will be needed to assess these properties in the context of extremes of variability to which they are exposed.

The final chapter, "Models and Management", brings together methods of synthesizing and using ecosystem studies in the coastal zone. The author discusses several computer simulations which attempt to reconstruct the system numerically from its physical, chemical and biological components. The general conclusion is that these exercises give valuable insights, but are far from being a predictive tool for environmental managers. Mann points out that the managerially useful models, such as those for biological oxygen demand in estuaries, would be considered by many researchers as unrealistic simplifications. The broad question he raises is whether simpler holistic models without explicit formulation of organism behaviour are not merely of empirical value, but a recognition of a general feature of the hierarchies within and between

ecosystems. This question is not new, nor confined to the coastal area, but it is central to both the future research and management philosophy for this part of our environment.

Ecology of Coastal Waters combines sound descriptive chapters with proper speculative flavouring. It is aimed at advanced students and professionals, and will be useful and interesting to all concerned with the edges of the sea. □

John H. Steele is Director of the Woods Hole Oceanographic Institution, Massachusetts.

Galileo rules, OK?

C. J. S. Clarke

Retarded Action-at-a-Distance: The Change of Force with Motion. By G. Burniston Brown. Pp.145. ISBN 0-904378-14-4. (Cortney Publications, 95-115 Windmill Road, Luton, UK: 1982.) £9.95.

IT IS in the nature of science to question, to refuse to treat anything as unalterable dogma; and so there is nothing improper in questioning the correctness of Maxwellian electrodynamics, special relativity and general relativity, as does the author of this book. In their place he proposes Euclidean space and time, with Galilean relativity, combined with retarded inter-particle action at a distance using unconventional force-laws. It is a polemical book, and so one looks for a clear critique of the central philosophical and experimental points where he differs from conventional wisdom. But one looks in vain: the experimental evidence is ignored, the philosophical arguments evaded.

For example, in his theory electromagnetic interactions are linear and linked to the Euclidean geometrical relations between the particles involved; thus it is crucial that a null result be obtained for the "gravitational bending of light" experiment. And so it is astonishing that his only reference on this subject is to a 1930 article criticizing the "early claims that the effect of gravitation on light at eclipses had been proved"; as if the author were unaware of the work of Counselman, Fomalont and Sramek that improved on the early eclipse observations by an order of magnitude in accuracy.

Gaia anew

J.E. Lovelock's *Gaia: A New Look at Life on Earth* (reviewed in *Nature* 282, 154; 1979) has now appeared in a paperback edition. The book is published by Oxford University Press, price £2.95.

The crucial difficulty with any direct action theory is the explanation of radiation-reaction, as manifested in the behaviour of charged particles in accelerators. It was this difficulty that led Wheeler and Feynman to propose their time-symmetric direct action theory, explicitly rejected by the author. Yet he concerns himself almost exclusively with phenomena such as current electricity, where it is possible to avoid the problem, and nowhere gives a clear account of radiation-reaction on individual particles.

The treatment of philosophical issues is no more satisfactory. The author is proposing a Galilean structure for space-time, but without any serious philosophical justification whatever. One supposes that he holds to the Kantian view that the Euclidean nature of space is part of its *a priori* status; but Kant was writing before Poincaré, Reichenbach and Gruenbaum, and before the Eötvös-type experiments of Dicke and Braginsky on which the relativistic concept of space is nowadays based. Dr Burniston Brown is writing after these writers and experimentalists have shown that it is far from obvious that the "space" used by astronomers is necessarily Euclidean. In 1880 this book might have been a useful contribution. In 1982 it is conspicuous only by its omissions. □

C. J. S. Clarke is a Lecturer in Mathematics at the University of York.

"THE MOST COMPLETE HISTORY OF THE ATOM TO BE PUBLISHED" *Energy Daily*

The Atomic Complex

by Bertrand Goldschmidt is an accurate, complete and fascinating worldwide political history and personal memoir of nuclear energy — from the development of the bomb in World War II to today's nuclear energy complex and proliferation problems.

Goldschmidt, a leading French scientist turned international statesman, reviews half a century of political moves, counter-moves, international intrigue and manipulation. *The Atomic Complex* is a pragmatic look at the nuclear world today, carefully examining the terror nuclear weapons represent, but at the same time stressing the benefits of nuclear energy.

Order this 500 page best seller today . . . "superbly done and reads easily."

**\$31 Hardbound
\$24 Softbound**

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Circle No. 12 on Reader Service Card.

BOOKS RECEIVED

Applied Biological Sciences

MITCHELL, M.S. and OETTGEN, H.F. (eds). *Hybridomas in Cancer Diagnosis and Treatment*. Progress in Cancer Research and Therapy, Vol.21. Pp.282. ISBN 0-89004-768-5. (Raven: 1982.) \$40.80.

NAJJAR, V.A. (ed.). *Immunologically Active Peptides*. Developments in Molecular and Cellular Biochemistry, Vol.2. Pp.135. ISBN 90-6193-842-2. (Junk, The Netherlands: 1981.) \$55, Dfl.125.

PALTI, J. *Cultural Practices and Infectious Crop Diseases*. Advanced Series in Agricultural Sciences, Vol.9. Pp.243. ISBN 3-540-11047-X/0-387-11047-X. (Springer-Verlag: 1981.) \$39.80.

PONTIFEX, G. (ed.). *Lung Cancer*. Proc. of the 1st European Symposium held September 7-13, 1980, Chalkidiki. Pp.381. ISBN 90-219-0499-3/0-444-90224-4. (Excerpta Medica/Elsevier: 1981.) \$90.75, Dfl.195.

ROBINSON, R. *Genetics for Dog Breeders*. Pp.264. ISBN 0-08-025917-0. (Pergamon: 1982.) £9.95, \$19.95.

RYMAN, N. (ed.). *Fish Gene Pools*. Preservation of Genetic Resources in Relation to Wild Fish Stocks. Proc. of an International Symposium held January 23-25, 1980, Stockholm. Ecological Bulletins, 34. Pp.111. ISBN 91-546-0299-8. (Swedish Council for Planning and Coordination of Research: 1981.) Np.

SALEM, M.R. (ed.). *Pediatric Anesthesia: Current Practice*, Vol.1. Pp.158. ISBN 0-12-615201-2. (Academic: 1981.) \$27.50.

SCHULTZ, V. and WARD WHICKER, F. *Radioecological Techniques*. Pp.298. ISBN 0-306-40797-3. (Plenum: 1982.) Np.

A SCIENTIFIC AMERICAN BOOK. *Industrial Microbiology and the Advent of Genetic Engineering*. Pp.108. Hbk ISBN 0-7167-1385-3; pbk 0-7167-1386-1. (Freeman: 1981.) Np.

SORENSEN, J.R., SWAZEY, J.P. and SCOTCH, N.A. *Reproductive Pasts Reproductive Futures: Genetic Counselling and Its Effectiveness*. Birth Defects: Original Article Series, Vol.XVII, 4. Pp.214. ISBN 0-8451-1044-6. (Alan R. Liss: 1981.) £21, DM 96.

SUTHERLAND, R.L. and JORDAN, V.C. (eds). *Non-steroidal Antioestrogens*. Molecular Pharmacology and Antitumour Activity. Pp.496. ISBN 0-12-677880-9. (Academic: 1982.)

SZARA, S.I. and LUDFORD, J.P. (eds). *Benzodiazepines: A Review of Research Results, 1980*. NIDA Research Monograph, 33. Pp.101. ISBN 017-024-01108-8. (DHHS, Rockville: 1981.) \$4.75.

THOMPSON, T. and JOHANSON, C.E. (eds). *Behavioral Pharmacology of Human Drug Dependence*. NIDA Research Monograph, 37. Pp.294. ISBN 017-024-01109-6. (DHHS, Rockville: 1981.) \$6.50.

VAN DEN BOSCH, R., MESSENGER, P.S. and GUTIERREZ, A.P. *An Introduction to Biological Control*. Pp.247. ISBN 0-306-40706-X. (Plenum: 1982.) Np.

WEISSMANN, G. (ed.). *Advances in Inflammation Research*, Vol.4. Pp.208. ISBN 0-89004-669-7. (Raven: 1982.) \$28.56.

YOSHIDA, H., HAGIHARA, Y. and EBASHI, S. (eds). *Biochemical Immunological Pharmacology*, Vol.4. Pp.280. ISBN 0-08-028024-2. (Pergamon: 1982.) Np.

YOSHIDA, H., HAGIHARA, Y. and EBASHI, S. (eds). *Cardio-renal and Cell Pharmacology*, Vol.3. Pp.340. ISBN 0-08-028023-4. (Pergamon: 1982.) Np.

YOSHIDA, H., HAGIHARA, Y. and EBASHI, S. (eds). *Clinical Pharmacology*. Teaching in Pharmacology, Vol. 6. Pp.315. ISBN 0-08-028026-9. (Pergamon: 1982.) Np.

YOSHIDA, H., HAGIHARA, Y. and EBASHI, S. (eds). *Toxicology and Experimental Models*, Vol. 5. Pp.355. ISBN 0-08-028025-0. (Pergamon: 1982.) Np.

Psychology

COLES, E.M. *Clinical Psychopathology: An Introduction*. Pp.502. Hbk ISBN 0-7100-0864-3; pbk ISBN 0-7100-0867-8. (Routledge & Kegan Paul: 1982.) Hbk £15; pbk £8.95.

OLSEN, P. (ed.). *Comprehensive Psychotherapy*, Vol. 3. Pp.156. ISBN 0-677-16369-A. (Gordon & Breach: 1981.) \$45.

SOLOMON, J. (ed.). *Alcoholism and Clinical Psychiatry*. Pp.238. ISBN 0-306-40794-9. (Plenum: 1982.) Np.

USDIN, E. (ed.). *Clinical Pharmacology in Psychiatry*. Pp.352. ISBN 0-444-00556-0. (Elsevier: 1981.) \$59, Dfl.135.

WOIBERG, L.R. *The Practice of Psychotherapy*. 506 Questions and Answers. Pp.416. ISBN 0-87630-290-8. (Brunner/Mazel, New York: 1982.) \$30.

Sociology

AULIN, A. *The Cybernetic Laws of Social Progress: Towards a Critical Social Philosophy and a Criticism of Marxism*. Pp.218. ISBN 0-08-025782-8. (Pergamon: 1981.) £22.50, \$45.

BLALOCK, A.B. and BLALOCK, H.M. Jr. *Introduction Social Research*, 2nd Edn. Pp.153. Hbk ISBN 0-13-496810-7; pbk ISBN 0-13-496802-6. (Prentice Hall: 1982.) Hbk £8.95; pbk £5.95.

EVANS, B. and WAITES, B. *IQ and Mental Testing: An Unnatural Science and its Social History*. Pp.228. Hbk ISBN 0-333-25648-4; pbk ISBN 0-333-25649-2. (Macmillan Press, London & New York: 1981.) Hbk np; pbk £4.95.

KONEČNÍ, V.J. and EBBESEN, E.B. (eds). *The Criminal Justice System. A Social-Psychological Analysis*. Pp.418. Hbk ISBN 0-7167-1312-8; pbk ISBN 0-7167-1313-6. (Freeman: 1982.) Hbk np; pbk np.

REYNOLDS, P.D. *Ethics and Social Science Research*. Pp.191. ISBN 0-13-290965-0. (Prentice Hall: 1982.) £7.45.

Anthropology

STEIN, P.L. and ROWE, B.M. *Physical Anthropology*, 3rd Edn. Pp.482. ISBN 0-07-061151-3. (McGraw-Hill: 1982.) \$18.95.

WOOLFSON, C. *The Labour Theory of Culture. A Re-examination of Engels' Theory of Human Origins*. Pp.124. ISBN 0-7100-0997-6. (Routledge & Kegan Paul: 1982.) £4.95.

General

CHARALAMBOUS, G. and INGLET, G. (eds). *The Quality of Foods and Beverages*. Chemistry and Technology, Vol.2. Proceedings of a Symposium of the 2nd International Flavor Conference, July 1981, Athens, Greece. Pp.383. ISBN 0-12-169102-0. (Academic: 1981.) \$26.50.

CHIGIER, N.A. (ed.). *Progress in Energy and Combustion Science*, Vol.6. Pp.395. ISBN 0-08-027153-7. (Pergamon: 1981.) £51, \$118.

COOPER, W.E. and SORENSEN, J.M. *Fundamental Frequency in Sentence Production*. Pp.213. ISBN 3-540-90510-3. (Springer-Verlag: 1981.) DM 48, \$22.90.

McMAHAN, J. *British Nuclear Weapons. For and Against*. Pp.165. Hbk ISBN 0-86245-047-0; pbk ISBN 0-86245-049-7. (Junction Books, London: 1981.) Hbk £9.95; pbk £3.95.

MEEKS, L. B. and HEIT, P. *Human Sexuality: Making Responsible Decisions*. Pp.503. ISBN 0-03-058366-7. (Saunders: 1982.) Np.

ORTNER, S. B. and WHITEHEAD, H. (eds). *Sexual Meanings. The Cultural Construction of Gender and Sexuality*. Pp.435. Hbk ISBN 0-521-23965-6; pbk ISBN 0-521-28375-2. (Cambridge University Press: 1982.) Hbk £25; pbk £7.95.

POLMAR, N. and ALLEN, T. B. *Rickover. Controversy and Genius: A Biography*. Pp.744. ISBN 0-671-24615-1. (Simon & Schuster, New York: 1982.) \$20.75.

RAYNOR, J. O. and ENTIN, E. E. *Motivation, Career Striving, and Aging*. Pp.406. ISBN 0-89116-189-9. (Hemisphere Publishing: 1982.) \$24.95.

RICHMOND, C.R., WALSH, P.J. and COPENHAVER, E.D. (eds). *Health Risk Analysis. Proceedings of the 3rd Life Sciences Symposium, Gatlinburg, Tennessee, October 1980*. Pp.483. ISBN 0-89168-042-X. (Franklin Institute Press, Pennsylvania: 1981.) \$40.

RIEPE, D. (ed.). *Asian Philosophy Today*. Pp.291. ISBN 0-677-15490-9. (Gordon & Breach: 1981.) \$24.

ROBERTS, M. J. and BLUHM, J. S. *The Choices of Power: Utilities Face the Environmental Challenge*. Pp.458. ISBN 0-674-12780-3. (Harvard University Press: 1981.) £17.50.

ROSS, S.D. *Learning and Discovery*. Pp.134. ISBN 0-677-05110-7. (Gordon & Breach: 1981.) \$29.50.

SALOMON, J.-J. *Prométhée Empêtré. La Résistance au Changement Technique*. Pp.174. Pbk ISBN 0-08-027064-6. (Pergamon: 1981.) Pbk np.

SAXENA, J. and FISHER, F. (eds). *Hazard Assessment of Chemicals*. Current Developments, Vol. 1. Pp.461. ISBN 0-12-312401-8. (Academic: 1981.) \$49.50.

SAYERS, J. *Biological Politics. Feminist and Anti-Feminist Perspectives*. Pp.235. Hbk ISBN 0-422-77870-2; pbk ISBN 0-422-77880-X. (Tavistock, London: 1982.) Hbk £10.95; pbk £4.95.

SHAFFER, M. *Life after Stress*. Pp.273. ISBN 0-306-40869-4. (Plenum: 1982.) \$15.95.

SHARPE, R.S. (ed.). *Research Techniques in Nondestructive Testing*, Vol. 5. Pp.332. ISBN 0-12-639055-X. (Academic: 1982.) £28.40, \$58.50.

SIGMON, B.A. and CYBULSKI, J.S. (eds). *Homo erectus Papers in Honor of Davidson Black*. Based on the Proceedings of an International Symposium, Cedar Glen, Ontario, October 1976. Pp.271. ISBN 0-8020-5511-7. (University of Toronto Press: 1982.) £21, \$37.50.

SOFER, S. S. and ZABORSKY, O. R. (eds). *Biomass Conversion Processes for Energy and Fuels*. Pp.420. ISBN 0-306-40663-2. (Plenum: 1981.) \$49.50.

STEPHENSON, R. M. *Living with Tomorrow: A Factual Look at America's Resources*. Pp.280. ISBN 0-471-09457-9. (Wiley: 1981.) £16.75, \$29.95.

SWERN, D. (ed.). *Bailey's Industrial Oil and Fat Products*, Vol. 2. 4th Edn. Pp.603. ISBN 0-471-83958-2. (Wiley: 1982.) £44, \$80.

UN ECONOMIC COMMISSION FOR EUROPE. *Protein and Energy Supply for High Production of Milk and Meat. A Symposium of the UN Economic Commission for Europe and the Food and Agriculture Organization*. Pp.191. ISBN 0-08-028909-6. (Pergamon: 1982.) £15, \$30.

WATERS, M. D. *et al.* (eds). *Short-Term Bioassays in the Analysis of Complex Environmental Mixtures II. Environmental Science Research*, Vol. 22. Proceedings of the 2nd Symposium held in Williamsburg, Virginia, March 1980. Pp.524. ISBN 0-306-40890-2. (Plenum: 1981.) \$59.50.

WATLING, R. *Three Bolbitaceae: Agrocycbe, Bolbitius & Conocycbe*. Pp.139. ISBN 0-11-491750-7. (HMSO: 1982.) £8.

WEATHERS, W. *The Broken Word. The Communication Pathos in Modern Literature. Communication and the Human Condition*, Vol. 1. Pp.226. ISBN 0-677-04990-0. (Gordon & Breach: 1981.) \$62.50.

WEBBER, B.L. and NILSSON, N.J. (eds). *Readings in Artificial Intelligence*. Pp.557. ISBN 0-935382-03-8. (W.H. Freeman: 1982.) £17.50.

WEISSMAN, R.F.E. *Ritual Brotherhood in Renaissance Florence: Population and Social Structure*. Pp.254. ISBN 0-12-744480-7. (Academic: 1982.) Np.

WILFORD, J. N. *The Mapmakers*. Pp.414. ISBN 0-86245-041-1. (Junction Books, London: 1981.) £9.95.

WINITZ, H. (ed.). *Native Language and Foreign Language Acquisition*, Vol. 379. Pp.378. Hbk ISBN 0-89766-147-8; pbk ISBN 0-89766-148-6. (New York Academy of Science: 1981.) Hbk \$75; pbk np.